

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 3, 2004, 06:16:43 ; Search time 80 Seconds
(without alignments)
8303.453 Million cell updates/sec

Title: US-10-063-730-115
Perfect score: 1197
Sequence: 1 cagcagtggtctctcagtc.....caaaaaaaaaaaaaaaaa 1197

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 10

Total number of hits satisfying chosen parameters: 257976

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCFUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1057	88.3	1380	4	US-09-620-312D-387
2	547	45.7	1428	4	US-09-620-312D-386
c 3	26	2.2	11050	4	US-10-204-708-86
c 4	25	2.1	87350	3	US-08-781-891-79
c 5	25	2.1	87350	4	US-09-618-166-79
c 6	25	2.1	87543	4	US-09-791-211-3
c 7	25	2.1	162450	4	US-09-345-882-1
8	24	2.0	414	1	US-08-377-687-48
9	24	2.0	414	1	US-08-771-192-48
10	24	2.0	414	3	US-08-971-982-48
11	24	2.0	414	4	US-09-077-951-19
12	24	2.0	414	4	US-09-077-948A-45
c 13	24	2.0	705	3	US-09-328-111-674
14	24	2.0	1396	1	US-08-123-161A-11
15	24	2.0	1396	1	US-08-483-278-11
16	24	2.0	1529	3	US-09-189-760-5
17	24	2.0	1529	3	US-09-188-811-5
18	24	2.0	1529	3	US-09-514-422-5
c 19	24	2.0	2017	1	US-07-667-276A-3
20	24	2.0	2041	4	US-09-149-476-131
21	24	2.0	2072	4	US-09-833-381-1791
22	24	2.0	2328	4	US-08-811-481-34
23	24	2.0	2328	4	US-09-876-527-34
24	24	2.0	2494	3	US-09-189-760-1
25	24	2.0	2494	3	US-09-514-422-1
26	24	2.0	13104	3	US-08-256-799-4
27	24	2.0	13104	3	US-08-462-437-4

28	23	1.9	55	2	US-08-771-624B-8	Sequence 8, Appli
29	23	1.9	104	4	US-09-621-976-9392	Sequence 9392, Ap
30	23	1.9	553	4	US-09-148-545-133	Sequence 133, App
31	23	1.9	557	4	US-09-148-545-80	Sequence 80, Appl
32	23	1.9	653	4	US-09-373-750-1	Sequence 1, Appli
c 33	23	1.9	976	1	US-08-125-628-7	Sequence 7, Appli
c 34	23	1.9	976	1	US-08-125-628-8	Sequence 8, Appli
35	23	1.9	990	4	US-09-461-325-94	Sequence 94, Appl
36	23	1.9	990	4	US-10-012-542-94	Sequence 94, Appl
37	23	1.9	1027	4	US-09-533-887-9	Sequence 9, Appli
38	23	1.9	1296	4	US-09-533-029-63	Sequence 63, Appl
39	23	1.9	1348	4	US-09-152-060-14	Sequence 14, Appl
40	23	1.9	1500	3	US-09-052-778-1	Sequence 1, Appli
41	23	1.9	1771	4	US-09-533-029-103	Sequence 103, App
42	23	1.9	1778	4	US-09-833-381-2044	Sequence 2044, Ap
43	23	1.9	2114	4	US-09-130-491-7	Sequence 7, Appli
c 44	23	1.9	2285	4	US-09-230-225B-1	Sequence 1, Appli
45	23	1.9	17949	3	US-09-087-465-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-620-312D-387
; Sequence 387, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 387
; LENGTH: 1380
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (131)..(1084)
US-09-620-312D-387

Query Match 88.3%; Score 1057; DB 4; Length 1380;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1107; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 73 TCACAGAAATGTGAAGACTGTCACTTTAAATGCAGAGCTTTTAAATCCAGAAAT 132
Db 145 TCACAGAAATGTGAAGACTGTCACTTTAAATGCAGAGCTTTTAAATCCAGAAAT 204

```
QY 133 ATGTAATCACTTAAGATTGTGGACTGGTGTGTTGGTATCTCTGGCCCTAACTCTAATTGT 192
Db 205 ATGTAATCACTTAAGATTGTGGACTGGTGTGTTGGTATCTCTGGCCCTAACTCTAATTGT 264
QY 193 CTTGTTTGGGGAGCAAGCACTTCTGGCGGGAGGTACCCAAAAGCCCTATGACATGGA 252
Db 265 CTTGTTTGGGGAGCAAGCACTTCTGGCGGGAGGTACCCAAAAGCCCTATGACATGGA 324
QY 253 GCACACTTTCTACAGCAATGGAGAGAGAGATTTCATGGAATTGATCCTGTGAC 312
Db 325 GCACACTTTCTACAGCAATGGAGAGAGAGATTTCATGGAATTGATCCTGTGAC 384
QY 313 CAGAACTGAAATATTCAGAAAGCGGAAATGSCACTGATGAAACATTTGGAAGTCACGACTT 372
Db 385 CAGAACTGAAATATTCAGAAAGCGGAAATGSCACTGATGAAACATTTGGAAGTCACGACTT 444
QY 373 TAAAAACGGATACACTGGGATCTACTTCTGGGTCTTCAAAATGTTTTATCAAACTCA 432
Db 445 TAAAAACGGATACACTGGGATCTACTTCTGGGTCTTCAAAATGTTTTATCAAACTCA 504
QY 433 GATTAAAGTGATTCCTGAATTTTCTGAAACAGAAAGAGAAATAGATGAGAAATCAAGAAAT 492
Db 505 GATTAAAGTGATTCCTGAATTTTCTGAAACAGAAAGAGAAATAGATGAGAAATCAAGAAAT 564
QY 493 TACCACAACTTTCTTTGAACAGTCAGTGATTTGGGTCCAGCAGCAAAAGCCTATTGAAA 552
Db 565 TACCACAACTTTCTTTGAACAGTCAGTGATTTGGGTCCAGCAGCAAAAGCCTATTGAAA 624
QY 553 CCGAGATTTCTTAAAAATCCAAATCTCGAGATTTGTGATAACGTGACCATGTATTG 612
Db 625 CCGAGATTTCTTAAAAATCCAAATCTCGAGATTTGTGATAACGTGACCATGTATTG 684
QY 613 GATCAATCCACTCTAATATCAGTTTCTGAGTTTACAAGACTTTGAGGAGGAGGAGAAGA 672
Db 685 GATCAATCCACTCTAATATCAGTTTCTGAGTTTACAAGACTTTGAGGAGGAGGAGAAGA 744
QY 673 TCTTCACTTTCTTCGCAACGAAAAAAGGATTTGAACAAAAATGAACAGTGGGTGTGCC 732
Db 745 TCTTCACTTTCTTCGCAACGAAAAAAGGATTTGAACAAAAATGAACAGTGGGTGTGCC 804
QY 733 TCAAGTGAAAGTAGAAGACCGTCCAGCCAGACGAAAGTGAGGAGGAGAACTTCCAAT 792
Db 805 TCAAGTGAAAGTAGAAGACCGTCCAGCCAGACGAAAGTGAGGAGGAGAACTTCCAAT 864
QY 793 AAATGACTACTAGAAAAAGAAATAGAAATTTGATCCCATGCTGGATCAGAGAGTTATTG 852
Db 865 AAATGACTACTAGAAATGGAATAGAAATTTGATCCCATGCTGGATCAGAGAGTTATTG 924
QY 853 TTGTAATTTACTGCGGTGAGGCAACCGCTATTGCCGCGCGTCTGTGAACCTTTACTAGG 912
Db 925 TTGTAATTTACTGCGGTGAGGCAACCGCTATTGCCGCGCGTCTGTGAACCTTTACTAGG 984
QY 913 CTACTACCCATATCCATCTACTTACCAGGAGGACGAGTCATCTGTCGTCTATCATGCC 972
Db 985 CTACTACCCATATCCATCTACTTACCAGGAGGACGAGTCATCTGTCGTCTATCATGCC 1044
QY 973 TTGTAATCTGGTGGGTGGCCGATGCTGGGGAGGGTCTAATAGGAGGTTTGAGCTCAAAAT 1032
Db 1045 TTGTAATCTGGTGGGTGGCCGATGCTGGGGAGGGTCTAATAGGAGGTTTGAGCTCAAAAT 1104
QY 1033 GCTTAACTGCTGGCAACATAATATAATGATGCTATTCAATGAATTTCTGCTATGAG 1092
Db 1105 GCTTAACTGCTGGCAACATAATATAATGATGCTATTCAATGAATTTCTGCTATGAG 1164
QY 1093 GCATCTGGCCCTGGTAGCCAGCTCTCCAGAAATTAATTTGATGAGTAATTTCTCTCTCATG 1152
Db 1165 GCATCTGGCCCTGGTAGCCAGCTCTCCAGAAATTAATTTGATGAGTAATTTCTCTCTCATG 1224
QY 1153 TTCTAATAAACTTTCTACATTATCACCA 1180
Db 1225 TTCTAATAAACTTTCTACATTATCACCA 1252
```

```
RESULT 2
US-09-620-312D-386
; Sequence 386, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyang
; APPLICANT: Chen, Rui-Hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 386
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (131)...(1132)
US-09-620-312D-386

Query Match 45.7%; Score 547; DB 4; Length 1428;
Best Local Similarity 100.0%; Pred. No. 8.8e-240;
Matches 547; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 634 AGTTCTGAGTTACAGACTTTGAGGAGGAGGAGAGATCTTCACCTTCTCCCAACGA 693
Db 754 AGTTCTGAGTTACAGACTTTGAGGAGGAGGAGAGATCTTCACCTTCTCCCAACGA 813
QY 694 AAAAAAGGATGAAACAAATGAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAGAC 753
Db 814 AAAAAAGGATGAAACAAATGAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAGAC 873
QY 754 CCGTCACGCAGACAAAGCAAGTGAGGAAGAACTTCCAAATAAATGACTATCTGAAATGG 813
Db 874 CCGTCACGCAGACAAAGCAAGTGAGGAAGAACTTCCAAATAAATGACTATCTGAAATGG 933
QY 814 AATAGAAATTTGATCCCATCTGATCGATGAGAGAGTTATTGTTGTTATTCTGCCGTGAGG 873
Db 934 AATAGAAATTTGATCCCATCTGATCGATGAGAGAGTTATTGTTGTTATTCTGCCGTGAGG 933
QY 874 CAACCGCTATTGCCGCGGTCTGTGAACCTTTACTAGGCTACTACCCATATCCATCTG 933
Db 994 CAACCGCTATTGCCGCGGTCTGTGAACCTTTACTAGGCTACTACCCATATCCATCTG 1053
QY 934 CTACCAAGGAGACGAGTCATCTGTCGTATCATGCTTGTAACTGTGGTGGGCGCG 993
Db 1054 CTACCAAGGAGACGAGTCATCTGTCGTATCATGCTTGTAACTGTGGTGGGCGCG 1113
QY 994 CATGCTGGGAGGGTCTAATAGAGGTTTGAAGCTCAAAATGCTTAAATCTGCTGCAACATA 1053
Db 1114 CATGCTGGGAGGGTCTAATAGAGGTTTGAAGCTCAAAATGCTTAAATCTGCTGCAACATA 1173
```

QY 1054 TAATAAATGATCTATTCAATGAATTTCTGCTATGAGGCATCTGGCCCTGGTAGCCA 1113
Db 1174 TAATAAATGATCTATTCAATGAATTTCTGCTATGAGGCATCTGGCCCTGGTAGCCA 1233
QY 1114 GCCTCCAGAAATTAAGTGTAGGTAATCCCTCTCTTCATGTCTTAATAAACTTCTACATTA 1173
Db 1234 GCCTCCAGAAATTAAGTGTAGGTAATCCCTCTCTTCATGTCTTAATAAACTTCTACATTA 1293
QY 1174 TCACCAA 1180
Db 1294 TCACCAA 1300

RESULT 3
US-10-204-708-86/c
; Sequence 86, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 86
; LENGTH: 11050
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-86

Query Match 2.2%; Score 26; DB 4; Length 11050;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1171 TTATCACCACCAAAAAAAAAAAAAAAAAAAAA 1196
Db 744 TTATCACCACCAAAAAAAAAAAAAAAAAAAAA 719

RESULT 4
US-08-781-891-79/c
; Sequence 79, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington

COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 87350 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-79
Query Match 2.1%; Score 25; DB 3; Length 87350;
Best Local Similarity 100.0%; Pred. No. 0.048;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1172 TATCACCACCAAAAAAAAAAAAAAAAAAAAA 1196
Db 29899 TATCACCACCAAAAAAAAAAAAAAAAAAAAA 29875
RESULT 5
US-09-618-166-79/c
; Sequence 79, Application US/09618166
; Patent No. 6583112
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/618,166
FILING DATE: 17-Jul-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McMasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 240052.419C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:


```
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
```

```
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:

Query Match          2.1%; Score 25; DB 4; Length 162450;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1173 ATCACCACCAAAAAAAAAAAAAAAAAAAAAA 1197
Db 151927 ATCACCACCAAAAAAAAAAAAAAAAAAAAAA 151903

RESULT 8
US-08-377-687-48
; Sequence 48, Application US/08377687
; Patent No. 5538525
; GENERAL INFORMATION:
; APPLICANT: BROEKAERT, WILLEM F.
; APPLICANT: CAMMUE, BRUNO P.A.
; APPLICANT: OSBORN, RUPERT W.
; APPLICANT: REES, SARAH B.
; APPLICANT: TERRAS, FRANKY R.G.
; APPLICANT: VANDERLEIDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/377,687
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/002,480
; FILING DATE: 04-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
```

```

;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..255
;
US-08-377-687-48
;
;
; 2.0%; Score 24; DB 1; Length 414;
; Best Local Similarity 100.0%; Pred. No. 0.16;
; Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 1174 TCACCAAAAAAAAAAAAAAAAAAAAA 1197
| | | | | | | | | | | | | | | | | |
Db 391 TCACCAAAAAAAAAAAAAAAAAAAAA 414
| | | | | | | | | | | | | | | | | |

RESULT 9
US-08-777-192-48
; Sequence 48, Application US/08777192
; Patent No. 5824869
; GENERAL INFORMATION:
; APPLICANT: BROEKAERT, WILLEM F.
; APPLICANT: CAMMUE, BRUNO P.A.
; APPLICANT: OSBORN, RUPERT W.
; APPLICANT: REES, SARAH B.
; APPLICANT: TERRAS, FRANKY R.G.
; APPLICANT: VANDERLEYDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: CUSHMAN DARB & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/777,192
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..255
;
US-08-777-192-48
;
;
; 2.0%; Score 24; DB 1; Length 414;
; Best Local Similarity 100.0%; Pred. No. 0.16;
; Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 1174 TCACCAAAAAAAAAAAAAAAAAAAAA 1197
| | | | | | | | | | | | | | | | | |
Db 391 TCACCAAAAAAAAAAAAAAAAAAAAA 414
| | | | | | | | | | | | | | | | | |

RESULT 11
US-09-077-951-19
; Sequence 19, Application US/09077951
; Patent No. 6372888
; GENERAL INFORMATION:
; APPLICANT: De Samblaux, Geneveva
; APPLICANT: Broekaert, Willem
; APPLICANT: Rees, Sarah
; TITLE OF INVENTION: Antifungal Proteins
; FILE REFERENCE: PPD50093
; CURRENT APPLICATION NUMBER: US/09/077,951
;

```

```
; CURRENT FILING DATE: 1999-03-11
; EARLIER APPLICATION NUMBER: GB 9525474.4
; EARLIER FILING DATE: 1995-12-13
; EARLIER APPLICATION NUMBER: PCT/GB96/03065
; EARLIER FILING DATE: 1996-12-12
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Raphanus sativus
US-09-077-951-19

Query Match          2.0%; Score 24; DB 4; Length 414;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1174 TCACCAAAAAAAAAAAAAAAAAAAAA 1197
Db 391 TCACCAAAAAAAAAAAAAAAAAAAAA 414

RESULT 12
US-09-077-948A-45
; Sequence 45, Application US/09077948A
; Patent No. 6605698
; GENERAL INFORMATION:
; APPLICANT: Van Amerongen, Aart
; APPLICANT: Fant, Franky
; APPLICANT: Borremans, Frans
; APPLICANT: De Samblaux, Genevieve
; APPLICANT: Sitjtsma, Lolke
; APPLICANT: Melloen, Robbert
; APPLICANT: Puijk, Wouter
; APPLICANT: Schaaper, Wilhelmus
; APPLICANT: Broekert, Willem
; APPLICANT: Van Gelder, Wilhelmus
; APPLICANT: Rees, Sarah
; TITLE OF INVENTION: Antifungal Proteins
; FILE REFERENCE: 109846-257(SYN-035)
; CURRENT APPLICATION NUMBER: US/09/077,948A
; CURRENT FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: PCT/GB96/03068
; PRIOR FILING DATE: 1996-12-12
; PRIOR APPLICATION NUMBER: GB 9606552.9
; PRIOR FILING DATE: 1996-03-28
; PRIOR APPLICATION NUMBER: GB 9525455.3
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Raphanus sativus
US-09-077-948A-45

Query Match          2.0%; Score 24; DB 4; Length 414;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1174 TCACCAAAAAAAAAAAAAAAAAAAAA 1197
Db 391 TCACCAAAAAAAAAAAAAAAAAAAAA 414

RESULT 13
US-09-328-111-674/c
; Sequence 674, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
```

```
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 674
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(705)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-674

Query Match          2.0%; Score 24; DB 3; Length 705;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1174 TCACCAAAAAAAAAAAAAAAAAAAAA 1197
Db 35 TCACCAAAAAAAAAAAAAAAAAAAAA 12

RESULT 14
US-08-123-161A-11
; Sequence 11, Application US/08123161A
; Patent No. 5449616
; GENERAL INFORMATION:
; APPLICANT: Campbell, Kevin P.
; APPLICANT: Roberts, Steven L.
; APPLICANT: Anderson, Richard D.
; APPLICANT: Ibraghimov, Oxana B.
; APPLICANT: Yang, Bin
; TITLE OF INVENTION: NUCLEIC ACID ENCODING DYSTROPHIN-ASSOCIATED
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kevin M. Farrell, P.C.
; STREET: P.O. Box 999
; CITY: York Harbor
; STATE: ME
; COUNTRY: USA
; ZIP: 03911
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/123,161A
; FILING DATE: 16-SEP-93
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/946,234
; FILING DATE: 14-SEP-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Farrell, Kevin M.
; REGISTRATION NUMBER: 35,505
; REFERENCE/DOCKET NUMBER: UIRF89-11A4
```

TELECOMMUNICATION INFORMATION:

TELEPHONE: (207) 363-0558

TELEFAX: (207) 363-0528

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 1396 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 4..1164

US-08-123-161A-11

Query Match

Best Local Similarity 2.0%; Score 24; DB 1; Length 1396;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1174 TCACCAAAAAAAAAAAAAAAAAAAAA 1197

|||||

Db 1366 TCACCAAAAAAAAAAAAAAAAAAAAA 1389

RESULT 15

US-08-483-278-11

; Sequence 11, Application US/08483278

; Patent No. 5686073

GENERAL INFORMATION:

; APPLICANT: Campbell, Kevin P.

; APPLICANT: Ibraghimov, Oxana B.

; APPLICANT: Ervasti, James M.

; APPLICANT: Leveille, Cynthia J.

; TITLE OF INVENTION: NUCLEIC ACID ENCODING DYSTROPHIN-ASSOCIATED

; TITLE OF INVENTION: PROTEIN

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Kevin M. Farrell, P.C.

; STREET: P.O. Box 999

; CITY: York Harbor

; STATE: ME

; COUNTRY: USA

; ZIP: 03911

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/483,278

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/123,161

; FILING DATE: 16-SEP-93

; ATTORNEY/AGENT INFORMATION:

; NAME: Farrell, Kevin M.

; REGISTRATION NUMBER: 35,505

; REFERENCE/DOCKET NUMBER: UIRF89-11A5

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (207) 363-0558

; TELEFAX: (207) 363-0528

; INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

; LENGTH: 1396 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: CDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 4..1164

; US-08-483-278-11

Query Match 2.0%; Score 24; DB 1; Length 1396;

Best Local Similarity 100.0%; Pred. No. 0.15;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1174 TCACCAAAAAAAAAAAAAAAAAAAAA 1197

|||||

Db 1366 TCACCAAAAAAAAAAAAAAAAAAAAA 1389

Search completed: September 3, 2004, 07:45:39

Job time : 82 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 3, 2004, 04:36:32 ; Search time 80 seconds
(without alignments)
8303.453 Million cell updates/sec

Title: US-10-063-730-115
Perfect score: 1197

Sequence: 1 cagcagtggtctctcagtc.....caaaaaaaaaaaaaaaaa l197

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1174	98.1	1380	4	US-09-620-312D-387
2	1116	93.2	1428	4	US-09-620-312D-386
3	1111	9.3	1006	1	US-08-241-465B-3
4	109.4	9.1	1006	1	US-08-241-465B-2
5	106.2	8.9	1006	1	US-08-241-465B-6
6	104.6	8.7	1006	1	US-08-241-465B-5
7	90.8	7.6	892	1	US-08-241-465B-4
8	90	7.5	363	1	US-08-047-033-6
9	90	7.5	364	1	US-08-047-033-5
10	90	7.5	365	1	US-08-047-033-4
11	90	7.5	366	1	US-08-047-033-3
12	86	7.2	892	1	US-08-241-465B-7
13	53.2	4.4	7218	1	US-08-232-463-14
14	49.2	4.1	230	1	US-08-047-033-7
15	41.8	3.5	12886	4	US-09-453-702B-14
16	41	3.4	505	4	US-09-621-976-15639
17	39	3.3	1664976	4	US-08-916-421B-1
18	36.4	3.0	323	4	US-09-621-976-10374
19	36	3.0	1465	4	US-09-573-906-1
20	35.8	3.0	246	4	US-09-134-001C-2521
21	35.6	3.0	1506	4	US-09-134-001C-1278
22	35.2	2.9	29485	4	US-09-785-381-6
23	35	2.9	5183	1	US-08-459-568-3
24	35	2.9	5183	2	US-08-399-411-3
25	35	2.9	5868	3	US-08-516-859A-3
26	35	2.9	5868	3	US-09-586-472-3
27	35	2.9	5868	4	US-09-528-706-3

c	28	34.8	2.9	5562	4	US-10-204-708-64	Sequence 64, Appl
	29	34.8	2.9	1664976	4	US-08-916-421B-1	Sequence 1, Appl
	30	34.6	2.9	2241	4	US-09-023-942A-9	Sequence 9, Appl
	31	34.4	2.9	1176	3	US-09-072-384-16	Sequence 16, Appl
	32	34.2	2.9	289	3	US-09-007-005-17	Sequence 17, Appl
	33	34.2	2.9	289	3	US-09-244-796-17	Sequence 17, Appl
	34	34	2.8	1335	4	US-09-543-681A-1183	Sequence 1183, Ap
	35	34	2.8	1686	4	US-09-134-001C-1354	Sequence 1354, Ap
	36	34	2.8	2065	3	US-09-370-473-5	Sequence 5, Appl
	37	33.8	2.8	832	4	US-09-134-000C-45	Sequence 45, Appl
c	38	33.8	2.8	832	4	US-09-621-976-2813	Sequence 2813, Ap
	39	33.8	2.8	1754	4	US-09-604-978-1	Sequence 1, Appl
	40	33.8	2.8	1754	4	US-09-604-728-1	Sequence 1, Appl
	41	33.6	2.8	1611	6	5213972-6	Patent No. 5213972
	42	33.6	2.8	3300	4	US-09-268-347-31	Sequence 31, Appl
c	43	33.6	2.8	8607	4	US-10-204-708-72	Sequence 72, Appl
	44	33.4	2.8	148	4	US-09-621-976-8100	Sequence 8100, Ap
	45	33.4	2.8	565	4	US-09-589-733C-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-620-312D-387
; Sequence 387, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 387
; LENGTH: 1380
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (131)..(1084)
US-09-620-312D-387

Query Match 98.1%; Score 1174; DB 4; Length 1380;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1177; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 CAGCAGTGTCCTCAGTCCTCTCAAAGCAAGAGTACTGTGTGAGAGACCAT 60
DB 73 CAGCAGGGGTCTCTCAGTCCTCTCAAAGCAAGAGTACTGTGTGAGAGACCAT 132

QY 61 GCGAAGAAATCCTCCAGAGAAATGTGAAGACTGTGCACATTTCTAAATGAGAGACTTTTAA 120
DB 133 GCGAAGAAATGCTCCAGAGAAATGTGAAGACTGTGCACATTTCTAAATGAGAGACTTTTAA 192
QY 121 ATCCAAAGAAATATGTAATCACTTAAGATTGTGGACTGGTGTGTTGTAATCCTGGCCCT 180
DB 193 ATCCAAAGAAATATGTAATCACTTAAGATTGTGGACTGGTGTGTTGTAATCCTGGCCCT 252
QY 181 RACTCTAAATGTCCTGTTTGGGGGAGCAAGCACTTCTGGCGGAGGTACCCCAAAAAGC 240
DB 253 AACTCTAAATGTCCTGTTTGGGGGAGCAAGCACTTCTGGCGGAGGTACCCCAAAAAGC 312
QY 241 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAAAGAAATTTACATGGAAT 300
DB 313 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAAAGAAATTTACATGGAAT 372
QY 301 TGATCTGTGACAGAACTGAATATTCAGAGCGGAATGACACTGATGAACATTGGA 360
DB 373 TGATCTGTGACAGAACTGAATATTCAGAGCGGAATGACACTGATGAACATTGGA 432
QY 361 AGTGACGACTTTTAAACCGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTT 420
DB 433 AGTACAGACTTTTAAACCGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTT 492
QY 421 TATCAAACTCAGATTAAAGTGAATTCCTGAATTTCTGAACAGAGAGAAATAGATGA 480
DB 493 TATCAAACTCAGATTAAAGTGAATTCCTGAATTTCTGAACAGAGAGAAATAGATGA 552
QY 481 GAATGAGAAATTTACACAACCTTTCTTGAACAGTCACTGATTTGGGTCCAGCAGAAA 540
DB 553 GAATGAGAAATTTACACAACCTTTCTTGAACAGTCACTGATTTGGGTCCAGCAGAAA 612
QY 541 GCCTATTGAAACCGGAGATTTCTTAAATAATCCAAAATTCGAGATTTGTGATAACGT 600
DB 613 GCCTATTGAAACCGGAGATTTCTTAAATAATCCAAAATTCGAGATTTGTGATAACGT 672
QY 601 GACCATGTATGGATCAATCCCACTCTAATATCAGTTTCTGAGTTTCAAGACTTTGAGGA 660
DB 673 GACCATGTATGGATCAATCCCACTCTAATATCAGTTTCTGAGTTTCAAGACTTTGAGGA 732
QY 661 GAGGGAGAGATCTTCACTTTCTCCACGAGAAAGGATTTGAACAAATGAACA 720
DB 733 GAGGGAGAGATCTTCACTTTCTCCACGAGAAAGGATTTGAACAAATGAACA 792
QY 721 GTGGGTGTCCTCAAGTGAAGTAGAAGACCCGTCACGCCAGACAAGTAGTGAGGA 780
DB 793 GTGGGTGTCCTCAAGTGAAGTAGAAGACCCGTCACGCCAGACAAGTAGTGAGGA 852
QY 781 AGAATTTCCAATAAATGACTATCTGAAATGGAATAGAAATTTGATCCCATGCTGGATGA 840
DB 853 AGAATTTCCAATAAATGACTATCTGAAATGGAATAGAAATTTGATCCCATGCTGGATGA 912
QY 841 GAGAGGTATTGTGTTTACTGCGCTGAGGCAACCGCTATTGCCGCGCTCTGTGA 900
DB 913 GAGAGGTATTGTGTTTACTGCGCTGAGGCAACCGCTATTGCCGCGCTCTGTGA 972
QY 901 ACCTTTACTAGGCTACTACCCATATCCATCTGCTACCAAGGAGGAGTCACTGTGCG 960
DB 973 ACCTTTACTAGGCTACTACCCATATCCATCTGCTACCAAGGAGGAGTCACTGTGCG 1032
QY 961 TGTCACTATGCTTTGAACTGTGGTGGTGGCCCGATGCTGGGAGGCTTAATAGAGGT 1020
DB 1033 TGTCACTATGCTTTGAACTGTGGTGGTGGCCCGATGCTGGGAGGCTTAATAGAGGT 1092
QY 1021 TTGAGCTCAATGCTTAAATCTGTCGCAACATATATAATGATGCTATTCAATGAAT 1080
DB 1093 TTGAGCTCAATGCTTAAATCTGTCGCAACATATATAATGATGCTATTCAATGAAT 1152
QY 1081 TCTGCCTATGAGGCATCTGGCCCTGTGTAGCAGCTCTCCAGAAATTTACTTTGAGGTAAT 1140
DB 1153 TCTGCCTATGAGGCATCTGGCCCTGTGTAGCAGCTCTCCAGAAATTTACTTTGAGGTAAT 1212
QY 1141 CCTCTCTTCATGTTCTTAATAAATTTCTACATTATCACAAAA 1182

DB 1213 CCTCTCTTCTGTTCTTAATAAACTTTCTACATTATCACCAACA 1254
RESULT 2
US-09-620-312D-386
; Sequence 386, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 386
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (131)..(1132)
US-09-620-312D-386
Query Match 93.2%; Score 1116; DB 4; Length 1428;
Best Local Similarity 95.7%; Pred. No. 1.2e-311;
Matches 1177; Conservative 0; Mismatches 5; Indels 48; Gaps 1;
QY 1 CAGCAGTGTCTCTCAGTCTCTCTCAAGCAAGAAAGAGTACTGTGTCTGAGAGACCAT 60
DB 73 CAGCAGGGTCTCTCAGTCTCTCTCAAGCAAGAAAGAGTACTGTGTCTGAGAGACCAT 132
QY 61 GCGAAGAAATCCTCCAGAGAAATGTGAAGACTGTGCACATTTCTAAATGAGAGACTTTAA 120
DB 133 GCGAAGAAATGTCCTCAGAGAAATGTGAAGACTGTGCACATTTCTAAATGAGAGACTTTAA 192
QY 121 ATCCAAAGAAATATGTAATCACTTAAGATTGTGGACTGGTGTGTTGTAATCCTGGCCCT 180
DB 193 ATCCAAAGAAATATGTAATCACTTAAGATTGTGGACTGGTGTGTTGTAATCCTGGCCCT 252
QY 181 AACTCTAAATGTCCTGTTTGGGGGAGCAAGCACTTCTGGCGGAGGTACCCCAAAAAGC 240
DB 253 AACTCTAAATGTCCTGTTTGGGGGAGCAAGCACTTCTGGCGGAGGTACCCCAAAAAGC 312
QY 241 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAAAGAAATTTACATGGAAT 300
DB 313 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAAAGAAATTTACATGGAAT 372
QY 301 TGATCTGTGACAGAACTGAATATTTCAAGCGGAAATGGCACTCATGAAACATTGGA 360
DB 373 TGATCTGTGACAGAACTGAATATTTCAAGCGGAAATGGCACTCATGAAACATTGGA 432

Db 539 AAGGACAAAGCTTCTTG---AATCTAAGGTGTAGAACTCTGGGTGACCTTCTTATT 595
QY 608 TATTGGATCAATCCCACTTAATATCAGTTCTTGAGTTTACAAAGACTTTGAGGAGGAGGA 667
Db 596 TTCTGGCTTAAACCACTTATCCAAAGAAATCCAGAGGGAAGAGAGAGTGTGTAAGA 655
QY 668 GAAGATCTTCACTTTCTGCGCAACGAAAAAAGGATTTGAACAAAAATGAACAGTGGGTG 727
Db 656 AAAATTGTTCCAACTACCAAAAAAGACCACACAGTGGACCAAGGAGCAACCCAGGCGT 715
QY 728 GTCCCTCAAGTGAAGTAGAAGACCGT-----CAGCCAGACAAGCAAGTAGAGAA 781
Db 716 GGAAGACTGAATAATGAACACAGACCCAGTGTTCAGAGGACTCAAGCCTTCAATCCT 775
QY 782 GAACCTTCAATAAATGACTATATCTGAAATGAATAGAAATTTGATCCCATGCTGATGAG 841
Db 776 GATAATCTTATCATCAGCAGGAAGGGAAGCATGACATTCGACCTTAGACTGATCAC 835
QY 842 AGAGTTATTGTTGATTACTGCGTCGAGCAACCGCTATTGCGGCGGTCTGTGAA 901
Db 836 GAAGGAATCTGTTGTATAGATGTAGGCGGAGCTACACCCACTGCCAGAAAGATCTGTGAA 895
QY 902 CCTTTACTAGGCTACTACCATATCCATCTCTACCAAGGAGGAGTCACTCTGCT 961
Db 896 CCCCTGGGGGCTATTACCCATGGCTTATAATATCAAGGCTGCGCTTGGCGCTGCAGA 955
QY 962 GTCATCATGCTTGTAACTGTGGTGGGCGCGCATGCTGGGAGGCTTAA 1012
Db 956 GTCATCATGCCATGTAGTGTGGTGGGCGCGCATCTTGGGATGTTGTA 1006

RESULT 4

US-08-241-465B-2
; Sequence 2, Application US/08241465B
; Patent No. 5719125
; GENERAL INFORMATION:
; APPLICANT: Fujio SUZUKI
; APPLICANT: Yuji HIRAKI
; APPLICANT: Kazuhiro TAKAHASHI
; APPLICANT: Junko SUZUKI
; APPLICANT: Jun KONDO
; APPLICANT: Aseko KOHARA
; APPLICANT: Akiko MORI
; APPLICANT: Ei YAMADA
; TITLE OF INVENTION: HUMAN CHONDROMODULIN-I PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; COUNTRY: D.C.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/241,465B
; FILING DATE: May 11, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-8850
; TELEFAX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1006 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: double
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: Human being (Homo Sapiens)
; IMMEDIATE SOURCE:
; CLONE: phCM-13-6
; FEATURE:
; NAME/KEY: P CDS
; LOCATION: 2..1003
; IDENTIFICATION METHOD: E
US-08-241-465B-2
Query Match 9.1%; Score 109.4; DB 1; Length 1006;
Best Local Similarity 49.5%; Pred. No. 1.1e-21;
Matches 411; Conservative 0; Mismatches 396; Indels 24; Gaps 4;
QY 203 GGGAGCAGCACTTCTGCGCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTTTC 262
Db 179 GGGGCTTCTACTTCTGGAAGGGAGCGACAGTCACATTTACATGTCCATTACCACTG 238
QY 263 TACGCAATGGAGAGAAGAAGATTTCATGGAAATTTGATCCTGTGACCAAGAACTGAA 322
Db 239 AGTATCAATGGGAAACTACAGATGGGTCAATGGAAATAGACGCTGGGAAACAACTTGGAG 298
QY 323 ATATTCAAGCGGAAATGGCACTCATGAACATTTGGAAGTGCACGACTTTAAAAACGGA 382
Db 299 ACCTTTAAATGGGAAGTGGAGCTGAAGAAGCAATTGCAGTTTAATGATTTCCAGATGCG 358
QY 383 TACACTGGCATCTACTTCTGGGTCTTCAAAAAATGTTTATCAAAACTCAGATTTAAAGTG 442
Db 359 ATCAGAAATTCGTTTGTCTGGAGGAGAGTCTACATTTAAAGCGCAAGTGAAGCT 418
QY 443 ATT-----CCTGAATTTTCTGAACAGAAAGAGAAATAGATGAGAAATGAAGAA---- 490
Db 419 CGTATTCCTGAGGTGGCGCGCTGACCAAAACAGAGCATCTCTCCAAACTGGGAAGCAAG 478
QY 491 ---ATTACCAACACTTTCTTGAACAGTCAGTGATTTGGTCCCGACGAGAAAGCCTATT 547
Db 479 ATCATGCCAGTCAAAATATGAAGAAATTTCTTATCTGGGTGGCTGTAGATCAGCCTGTG 538
QY 548 GAAACCGGAGATTTCTTAAAAATTTCCAAAATTTCTGGAGATTTGTGATAACTGACCATG 607
Db 539 AAGGAACAGCTTCTTGAG---TTCTAAGGTGTAGAACTCTGCGGTGACCTTCTTATT 595
QY 608 TATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTTACAAGACTTTGAGGAGGAGGA 667
Db 596 TTCTGGCTTAAACCAACCTTATCCAAAAAGAAATCCAGAGGGAAGAGAGAGTGTGTAAGA 655
QY 668 GAAGATCTTCACTTTCTGCGCAACGAAAAAAGGATTTGAACAAAAATGAACAGTGGGTG 727
Db 656 AAAATTGTTCCAACTACCAAAAAAGACCACACAGTGGACCCAGGAGCAACCCAGGCGT 715
QY 728 GTCCCTCAAGTGAAGTAGAAGACCCGT-----CAGCCAGACAAGCAAGTAGAGAA 781
Db 716 GGAAGACTGAATAATGAACACAGACCCAGTGTTCAGAGGACTCACAAGCCTTCAATCCT 775
QY 782 GAACCTTCAATAAATGACTATATCTGAAATGGAATGGAATTTGATCCCATGCTGATGAG 841
Db 776 GATAATCTTATCATCAGCAGGAAGGGAAGCATGACATTCGACCTTAGACTGATCAC 835
QY 842 AGAGGTATTGTTGATTACTGCGCTCGAGGCAACCGCTATTTCGCGCGCTCTGTGAA 901
Db 836 GAAGGAATCTGTTGTATAGATGTAGCGGAGCTACACCATCTGCCAGAAATCTGTGAA 895
QY 902 CCTTTACTAGGCTACTACCCATATCCATCTCTACCAAGGAGGAGTCACTCTCTCGT 961
Db 896 CCCCTGGGGGCTATTACCCATGGCTTATAATATCAAGGCTGCGCTTGGCGCTGCAGA 955
QY 962 GTCATCATGCTTGTAACTGTGGTGGGCGCGCATGCTGGGAGGCTTAA 1012
Db 956 GTCATCATGCCATGTAGTGTGGTGGGCGCGCATCTTGGGATGTTGTA 1006

```
RESULT 5
US-08-241-465B-6
; Sequence 6, Application US/08241465B
; Patent No. 5719125
; GENERAL INFORMATION:
; APPLICANT: FUJIO SUZUKI
; APPLICANT: YUJI HIRAKI
; APPLICANT: KAZUHIRO TAKAHASHI
; APPLICANT: JUNKO SUZUKI
; APPLICANT: JUN KONDO
; APPLICANT: ATSUKO KOHARA
; APPLICANT: AKIKO MORI
; APPLICANT: EI YAMADA
; TITLE OF INVENTION: HUMAN CHONDROMODULIN-I PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; COUNTRY: D.C.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/241,465B
; FILING DATE: May 11, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1006 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: Human being (Homo Sapiens)
; FEATURE:
; NAME/KEY: P CDS
; LOCATION: 2..1003
; IDENTIFICATION METHOD: E
; US-08-241-465B-6

Query Match      8.9%; Score 106.2; DB 1; Length 1006;
Best Local Similarity 49.2%; Pred. No. 9e-21;
Matches 409; Conservative 0; Mismatches 398; Indels 24; Gaps 4;

QY 203 GGGAGCAAGACATCTTGGCCGGAGGTACCAAAAAAGCCCTATGACATGGAGCACACTTTC 262
    |||
DB 179 GGGGCCCTTCTACTCTTGGAGGGGAGGAGGACAGTCACATTTACATGTCCTTACACCATG 238
    |||

QY 263 TACAGCAATGGAGAGAGAGAGATTTACATGGAATTTGATCCTGTGACCAGAACTGAA 322
    |||
DB 239 AGTATCAATGGGAACTTACAAAGATGGGTCAATGGAATAGACGCTGGGAACTTGGAG 298
    |||

QY 323 ATATTGAGAGCGGAAATGCGCATGTGAAACATTTGGAAGTGCACGACTTTAAAAACGGA 382
    |||
DB 299 ACCTTTAAAAATGGGAAGTGGAGCTGGAAGAGCAATTGCAGTTAATGATTTCCAGAATGC 358
    |||

QY 383 TACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTATCAAACTCAGATTAAAGTG 442
    |||
DB 359 ATCAGAGAAATTCGTTTCTGGAGGAGAGAGTGCTACATTTAAAGCGCAAGTGAAGGCT 418
    |||
QY 443 ATT-----CCTGAATTTTCTGAACAGAGGAGAAATAGATGAGAAATGAAGAA----- 490
    |||

Db 419 CGTATTCTCTGAGGTGGCGCGGTGACCAAAACAGAGCATCTCTCTCCAACTGGRAAGCAAG 478
    |||
QY 491 ---ATTACCAAACTTTCTTTTGAACAGTCACTAGTATTTGGGTCCCGACAGAAAAGCCTATT 547
    |||
DB 479 ATCATGCCAGTCAAAATATGAAGAAAATTTCTTATCTGGGTGGCTGTAGATCAGCCTGTG 538
    |||
QY 548 GAAAACCGAGATTTTCTTAAAAATTCAAAATTTCTGAGATTTTGTGATAACGTGACCATG 607
    |||
DB 539 AAGGACAAACAGCTTCTTG---AATTTCTAAGGTGTAGAACTCTGCGGTGACCTTCTTATT 595
    |||
QY 608 TATTTGATCAATCCCACTCTTAATATCAGTTTCTGAGTTTACAAGACTTTTGAGGAGAGGGA 667
    |||
DB 596 TTCTGCTTAAACCAACCTATCCAAAAGAAATCCAGAGGGAAGAGAGAGTGTGTAAGA 655
    |||
QY 668 GAAGATCTTCACATTTCTTCCCAACGAAAAAAGGAGTTGAACAAAAATGAACAGTGGGTG 727
    |||
DB 656 AAAATTGTTTCCAACTACCAAAAAAGACCACACAGTGGACACCGGAGCAACCCAGGGGCT 715
    |||
QY 728 GTCCTCAAGTGAAGTAGAGAGACCCGT-----CAGCCAGACAAAGCAAGTGAGAA 781
    |||
DB 716 GGAAGACTGAATAATGAACACGACCAAGTGTTCAGAGGACTCAAGCCTTCAATCT 775
    |||
QY 782 GAATTTCCAATAAATGACTATCTGAAAATGGAATAGAAATTTGATCCCATGCTGGATGAG 841
    |||
DB 776 GATAATCCTTATCATCAGCAGGAGGAGGAAAGCATGACATTCGACCTAGACTGGATCAC 835
    |||
QY 842 AGAGGTATTGTTGTTTACTTCTCCGTCGAGGCAACCGCTATTGGCCGCCCGCTGTGAA 901
    |||
DB 836 GAAGGAATCTGTTGTATAGAAATGTAGCGGAGCTACACCCACTGCCAAGAGATCTGTGAA 895
    |||
QY 902 CCTTTACTGGCTACTACCCATATCCATATCTGTTACCAAGGAGGAGCAGTCACTGTGCT 961
    |||
DB 896 CCCTGGGGGCTATTACCCATGCGCTTATTAATATCAAGCTGCCCTTCGGCTTGCGA 955
    |||
QY 962 GTCATCATGCTTTGTAACCTGTTGGTGGCCCGCATGCTGGGGAGGGTCTAA 1012
    |||
DB 956 GTCATCATGCCATGTAGTGGTGGTGGCCCGCTATCTTGGGCATGTGTGA 1006
    |||

RESULT 6
US-08-241-465B-5
; Sequence 5, Application US/08241465B
; Patent No. 5719125
; GENERAL INFORMATION:
; APPLICANT: FUJIO SUZUKI
; APPLICANT: YUJI HIRAKI
; APPLICANT: KAZUHIRO TAKAHASHI
; APPLICANT: JUNKO SUZUKI
; APPLICANT: JUN KONDO
; APPLICANT: ATSUKO KOHARA
; APPLICANT: AKIKO MORI
; APPLICANT: EI YAMADA
; TITLE OF INVENTION: HUMAN CHONDROMODULIN-I PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; COUNTRY: D.C.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/241,465B
; FILING DATE: May 11, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1006 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: Human being (Homo Sapiens)
; FEATURE:
; NAME/KEY: P CDS
; LOCATION: 2..1003
; IDENTIFICATION METHOD: E
; US-08-241-465B-6
```

```
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1006 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: Human being (Homo Sapiens)
; FEATURE:
; NAME/KEY: P CDS
; LOCATION: 2..1003
; IDENTIFICATION METHOD: E
; US-08-241-465B-5

Query Match      8.7%; Score 104.6; DB 1; Length 1006;
Best Local Similarity 49.1%; Pred. No. 2.6e-20;
Matches 408; Conservative 0; Mismatches 399; Indels 24; Gaps 4;

QY 203 GGGAGCAAGCACTTCGTGGCCGAGGTACCCAAAAAGCCTATGACATGAGCAGCACACTTTC 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 179 GGGGCGCTTCTACTTCTGGAAGGGGAGCGACAGTCACATTTACAATGTCCATTACACCATG 238

QY 263 TACAGCATGGAGAGAGAGAGATTTACATGGAAATTCCTCTGACCGAAGCTGAA 322
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 239 AGTATCATGGGAACTACAAGATGGGTCAATGGAAATAGACGCTGGGAACAACCTTGGAG 298

QY 323 ATATTACAGAGCGGAAATGGCACTCATGAAACATTTGGAAGTGCACGACTTTAAAAACGGA 382
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 299 ACCTTTAAATGGGAATGGAGCTGAGAGCAATTTGCAGTTAATGATTTCCAGATGCG 358

QY 383 TACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTATCAAAACTCAGATTAAAGTG 442
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 359 ATCACAGGAATTCGTTTTCTGGAGGAGAGAGTGTCTACATTTAAAGCGCAAGTGAAGGT 418

QY 443 ATT-----CCTGAATTTCTGACCAAGAGAGAAATAGATGAGATGAGAA---- 490
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 419 CGTATTCCTGAGGTGGGCGGTGACCAACAGAGCATCTCTTCCAAACTGGAAGCAAG 478

QY 491 ---ATTACACAACTTCTTTGAAACAGTCAGTGAATTTGGTCCACGAGAAAAGCCTATT 547
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 479 ATCATGCCAGTCAATATCAAGAAATTCCTTATCTGGTGGCTGTAGATCAGCCTGTG 538

QY 548 GAAACCGAGATTTCTTAAAAATTTCCAAAATTTCTGGAGATTGTGATTAACGTGACCATG 607
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 539 AAGGACAAACAGCTTCTTGAG---TTCTAAGGTGTTAGAACTCTGCGGTGACCTTCTCTATT 595

QY 608 TATTGGATCAATCCCACTCTAATATATCAGTTTCTGAGTTACAAGACTTTCGAGGAGGGA 667
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 596 TTCTGGCTTAAACCAACCTATCCAAAAGAAATCCAGAGGGAAGAAGAGAGATGGTAAAGA 655

QY 668 GAAGATCTTCACTTTCCTGCCAACGAAAAAAGGGAATTGAACAAATGAACAGTGGGTG 727
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 656 AAAATTGTTTCCAACTACCAACAAAAAGACCACAGTGGACCGAGGAGCAACCCAGCGCT 715

QY 728 GTCCCTCAAGTGAAGTGAAGAACCCGT-----CAGCCNAGACAAGCAAGTGAAGAA 781
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 716 GGAAGACTGAATTAATGAACCAAGACCCAGTGTTCAGAGGAGACTTCAAAAGCTTCAATCCT 775

QY 782 GAACTTCCCAATATGACTATCTAGAAATGGAATAGAAATTTGATCCCATCTGATGATGAG 841
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 776 GATAATCCCTTATCATCAGCAGGAGGGAAGCATGACATTCGACCCCTAGACTGGATCAC 835

QY 842 AGAGGTATTGTGTATTACTGCGGTGAGGCAACCGCTATTGTGCGCCGCGTCTGTGAA 901
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 836 GAAGGAATCTGTGTATAGAAATGTAGGCGGAGCTACACCCACTGCCAGAGAAGATCTGTGAA 895

QY 902 CCTTTACTAGCTACTACCATATCCATCTGCTACCAAGGAGGAGAGTCACTCTGCT 961
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

QY 879 GCTATTGCCCGCGCTGTGTGAACCTTTACTAGGCTACTACCCATATCCATCTACTGCTACC 938
Db 759 CCCACTGCCAGAGATCTGTGAACCCCTGGGGGCTATTACCCATGGCCTTTATAATTATC 818
QY 939 AAGGAGCAGAGTCATCTGTGGTGTATCATGTGCTTTGTAACTGGTGGGCGCGCATGC 998
Db 819 AAGCTCGCCGTTCCGCTCGAGAGTCATCATGCCATGTAGCTGGTGGTGGCGCGCATCC 878
QY 999 TGGGAGGGTCTAA 1012
Db 879 TGGCATGGTGTA 892

RESULT 8

US-08-047-033-6
; Sequence 6, Application US/08047033
; Patent No. 5444157
; GENERAL INFORMATION:
; APPLICANT: FUJIO SUZUKI et al.
; TITLE OF INVENTION: CHONDROMODULIN-I PROTEIN
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/047,033
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/821,859
; FILING DATE:
; APPLICATION NUMBER: US/07/745,497
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 363 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: bovine
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE: fetal cartilage
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:

; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 249
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "N is T, G, A or C"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-08-047-033-6
Query Match 7.5%; Score 90; DB 1; Length 363;
Best Local Similarity 63.8%; Pred. No. 2.5e-16;
Matches 132; Conservative 2; Mismatches 73; Indels 0; Gaps 0;
QY 806 GAAATGGATAGAAATTGATCCCATGCTGGATGAGAGAGGTATTGTTGTATTACTGC 865
Db 157 GGAGAAAGCATGACATTCGACCCCGAGACTGGATCATGAGGAATCTGCTGTATAGATGC 216
QY 866 CGTCGAGGCAACCGCTATTGCGCGCGCTGNGAACCTTTACTAGGCTACTACCCATAT 925
Db 217 AGGAGAGCTACACCATCTGCCAGAGATCKNGAGCCTCTGGGGGCTACCCCATGG 276
QY 926 CCATCTGCTTACCAAGGAGGAGCATCTGTGCTGTCATCATGCTTGTAACTGTTGG 985
Db 277 CCCTATAACTACCAAGGCTGCCGCTCCGCTCGAGAGTCATCATGCCCTGTAGTGGTGG 336
QY 986 GTGGCCGCGCATGCTGGGAGGGTCTAA 1012
Db 337 GTGGCCGCGCATCTCTGGGCGCATGGTGTGA 363

RESULT 9

US-08-047-033-5
; Sequence 5, Application US/08047033
; Patent No. 5444157
; GENERAL INFORMATION:
; APPLICANT: FUJIO SUZUKI et al.
; TITLE OF INVENTION: CHONDROMODULIN-I PROTEIN
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/047,033
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/821,859
; FILING DATE:
; APPLICATION NUMBER: US/07/745,497
; FILING DATE:

```
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 364 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: modified site
LOCATION: 250
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "N is T, G, A or C"
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-047-033-5
Query Match 7.5%; Score 90; DB 1; Length 364;
Best Local Similarity 63.8%; Pred. No. 2.5e-16;
Matches 132; Conservative 2; Mismatches 73; Indels 0; Gaps 0;
QY 806 GAAATGGAATAGATTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTTACTGC 865
Db 158 GGAGAAAGCATGACATTCGACCCAGACTGGATCATGAAGGAATCTGCTATAGATGC 217
QY 866 CGTCGAGGCAACCGGTATTGCGCGCGCTGTGTAACCTTTTACTAGGCTACTACCCATAT 925
Db 218 AGGAGGAGCTACACCCACTGCCAGAGATCKNGAGCCTCTGGGGGGCTACCCCATGG 277
QY 926 CCATCTGCTACCAAGGAGACGAGTCATCTGCTGTCATCATGCTTGTAACTGGTGG 985
Db 278 CCTATAAATACCAAGGCTGCCGTTCCGCTGCAGAGTCATCATGCTTGTAGCTGGTGG 337
QY 986 GTGCCCCGATGCTGGGGGGGTCTAA 1012
Db 338 GTGCCCCGATCTCTGGGCATGGGTGA 364
```

RESULT 10
US-08-047-033-4
; Sequence 4, Application US/08047033
; Patent No. 544157
; GENERAL INFORMATION:
; APPLICANT: Fujio SUZUKI et al.
; TITLE OF INVENTION: CHONDROMODULIN-I PROTEIN
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/047,033
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/821,859
; FILING DATE:
; APPLICATION NUMBER: US/07/745,497
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: bovine
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE: fetal cartilage
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 251
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "N is T, G, A or C"
; PUBLICATION INFORMATION:
; AUTHORS:

```

; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
;
US-08-047-033-4

Query Match          7.5%; Score 90; DB 1; Length 365;
Best Local Similarity 63.8%; Pred. No. 2.5e-16;
Matches 132; Conservative 2; Mismatches 73; Indels 0; Gaps 0;

QY      806  GAAATGGGAATAGAAATTGATCCCATCTCGATGAGAGAGGTTATTGTTGTTACTGC 865
Db      159  GGAGAAAGCATGACATTCGACCCCGAGACTGATCATGAAGGAATCTGCTGTATAGAAATGC 218

QY      866  CGTCGAGGCAACCGCTATTTCGCCCGCGTCTGTGAACTTTACTAGGCTACTACCCATAT 925
Db      219  AGGAGGAGCTACACCCACTCCGAGAGCTCKNGAGCCTCTGGGGGGCTACACCCCATGG 278

QY      926  CCATCTGCTACCAAGGAGGACAGTCACTCTGCTGTCATCATGCTTTGTAACCTGGTGG 985
Db      279  CCCTATAACTACAGGGCTCCGCTTCGCGCTGCAGAGTCATCATGCCCTGTAGCTGGTG 338

QY      986  GTGCCCGCATGCTGGGAGGGCTAA 1012
Db      339  GTGCCCGCATCTCTGGGCATGGTGTGA 365

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	7.5%;	Score 90;	DB 1;	Length 366;
Best Local Similarity	63.9%;	Pred. No. 2.5e-16;		
Matches 132;	Conservative	2;	Mismatches 73;	Indels 0; Gaps 0
QY	806	GAATAATGGGAATGAATTTGATCCCATCTCGATGAGAGAGGTTATTGTTGTATTACTTACGC	865	
Db	160	GGGAAGCATGACATTCGACCCCGACTGGATCATGAAGGAATCTGCTGTATAGAAATGC	219	
QY	866	CGTCGAGGCAACCGCTATTTCGCGCGCGGTCTGTGAAACCTTTACTAGGCTACTACCCATAT	925	
Db	220	AGAGGAGGCTTACACCCACTGCCAGAATCKXGAGGCTCTGGGGGGGTACACCCCATGG	279	
QY	926	CCATCTGCTTACCAAGGAGGACGAGTCATCTGTGCTGCATCATGCTTGTAACTGGTGG	985	
Db	280	CCCTATTAATACCAAGGCTCGGTTCCGCTTCGAGAGTCATCATGCTCTGTAGCTGGTGG	339	
QY	986	GTGCGCCCGCATCTGGGAGGGTCTAA	1012	
Db	340	GTGGCCCGCATCTGGGCATGGTGTGA	366	

RESULT 12
US-08-241-465B-7
Sequence 7, Application US/08241465B
Patent No. 5719125
GENERAL INFORMATION:
APPLICANT: FUJIO SUZUKI
APPLICANT: YUJI HIRAKI
APPLICANT: KAZUHIRO TAKAHASHI
APPLICANT: JUNKO SUZUKI

APPLICANT: Jun KONDO
APPLICANT: Atsuko KOHARA
APPLICANT: Akiko MORI
APPLICANT: Ei YAMADA
TITLE OF INVENTION: HUMAN CHONDROMODULIN-I PROTEIN
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSER: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
COUNTRY: D.C.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/241,465B
FILING DATE: May 11, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 892 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: Human being (Homo Sapiens)
FEATURE:
NAME/KEY: P CDS
LOCATION: 2 .. 899
IDENTIFICATION METHOD: E
US-08-241-465B-7

Query Match 7.2%; Score 86; DB 1; Length 892;
Best Local Similarity 58.7%; Pred. No. 5.6e-15;
Matches 149; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 759 ACGCCAGACAGCAAGTGGAGGAGAACTTCCAAATAAGTACTATCTGAAATGGAATAG 818
Db 639 AGGACTCAAGCCCTTCAATCCTGATAATCTTATCATCAGAGGAGGGGAAAGCATGA 698
QY 819 AATTTGATCCCATGCTGGATGAGAGAGGTATTGTGTATTCTGCGCTCGAGGCAACC 878
Db 699 CATTGACCCCTAGATGGATCAGAGGAATCTGTTGATAGAAATGAGCGGAGCTACA 758
QY 879 GCTATTGGCGCGCTGTGTGAACCTTTACTAGGCTACTACCCATATCCATCTGCTACC 938
Db 759 CCACCTGCCAGAAGATCTGTGAACCCCTGGGGGGCTATTACCCATGCGCTTATAATTATC 818
QY 939 AAGGAGGAGCAGTCTCTCTGTGTCATCATCCTTTGTAACCTGTGGTGGGCCCGCATGC 998
Db 819 AAGGCTGCGGTTCGGCTGCAGAGCATCATCCATGATAGCTGGTGGGTCGCGCTATCT 878
QY 999 TGGGAGGAGGCTTAA 1012
Db 879 TGGGATGGTGTGA 892

RESULT 13

US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:

APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match 4.4%; Score 53.2; DB 1; Length 7218;
Best Local Similarity 4.1%; Pred. No. 4.5e-05;
Matches 16; Conservative 216; Mismatches 154; Indels 0; Gaps 0;

QY 435 TTAAAGTGATCTCCTGAATTTCTGAACCCAGAGAGAAATAGATGAGATGAAGAAATTA 494
Db 1459 TTAAAGAGATAGAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1400

QY 495 CCACAACITTTCTTGAACAGTCAGTGATTTGGGTCCAGCAGAAAGCCTATTGAAACC 554
Db 1399 RRR 1340

QY 555 GAGATTTCTTAAATAATCCAAAATCTGGAGATTGTGATAACCTGACCATGTATTGA 614
Db 1339 RRR 1280

QY 615 TCAATCCACTATATCAGTTTCTGAGTTACAAGACTTTGAGAGGAGGAGGAGATC 674
Db 1279 RRR 1220

QY 675 TTCCTTTCTGCCAACGAAAAAGGATTTGAACAAATGAACAGTGGTGGTCCCTC 734
Db 1219 RRR 1160

QY 735 AAGTGAAGTAGAGAAGACCCGCTACGCCAGACAGCAAGTAGGAGGAACTTCCAATAA 794
Db 1159 RRR 1100

QY 795 ATGACTACTGAAATGGAATAGAA 820

Db

PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:

US-08-047-033-7

Query Match 4.1%; Score 49.2; DB 1; Length 290;
Best Local Similarity 60.4%; Pred. No. 0.00013;
Matches 81; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 806 GAAATGGAATAGATTGTCCCATCTCGTGGATGAGAGAGTTATTCTTTACTTGTC 865
DB | ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 157 GGAGAAGCATTGCACATTCGCCACCAGACTGGATCATGAGGCACTCTGCTATAGATGC 216

QY 866 CGTCGAGGCAACCGCTATTGCGCGGCTCTGTGAACCTTTACTAGGCTACTACCCATAT 925
DB | ||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 217 AGGAGGAGCTACACCACTGCCAAGAGATCTGTGAGCCTCTGGGGGGCTTACCCCATG 276

QY 926 CCATACTGCTACCA 939
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 277 CCTATRACTACCA 290

RESULT 15

US-09-453-702B-14
Sequence 14, Application US/09453702B
Patent No. 6365723
GENERAL INFORMATION:
APPLICANT: Blattner, Frederick R.
 Burland, Valerie
 Perna, Nicole T.
 Plunkett, Guy
 Welch, Rod
TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 12886
TYPE: nucleic acid

```
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-453-702B-14

Query Match      3.5%; Score 41.8; DB 4; Length 12886;
Best Local Similarity 45.5%; Pred. No. 0.12;
Matches 135; Conservative 0; Mismatches 161; Indels 1; Gaps 1;

QY 415 ATGTTTTCATAAACTCAGATTAAAGTGATTCCTGAATTTCTGAACCGAGAGGAAAT 474
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DQ 3283 ATGAATTAAGAGTGTAACTGAGAGTCGTTAATCTTTCTGATATTCCTAGCGAGGAGA 3342
QY 475 AGATGAGATGAAGAAATTACCAAACTTTCTTTGAACAGTCAGTGATTTGGGTCCCAGC 534
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DQ 3343 CAAAAAGAAAGTCTGAATTTAAATTCCTTTTAACTCAGAGAGATGATCCTCTTAGC 3402
QY 535 AGAAAAGCCTATTGAAAACCGAGATTTCTTAAAAAATTCAAAAATTCGGAGATTTGTGA 594
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DQ 3403 AGAAAACCCCAATAAAAAAGCCAGATATTTTAAAAAAATC-CAATTGCCAATATTTATGA 3461
QY 595 TAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAGACTT 654
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DQ 3462 GGAAATTGAAAGNGCAATTTATTTATCTGAAATANGANCAGNCNNTGNTTCCANAGAACN 3521
QY 655 TGAGGAGGAGGGAGAGATCTTCACCTTCTGCGAACGAAAAAAGGGGATTGAACA 711
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DQ 3522 TNANAAGNANAGCCANATNTTNAATCNTCGAANANANATTNAAAGGNATTGNTNA 3578
```

Search completed: September 3, 2004, 06:18:17
Job time : 85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 3, 2004, 07:44:08 ; Search time 2370 Seconds
(without alignments)
15082.291 Million cell updates/sec

Title: US-10-063-730-115

Perfect score: 1197

Sequence: 1 cagcagtggtctctcagtc.....caaaaaaaaaaaaaaaaa 1197

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

```

EST.*
1: em_estba:*
2: em_estba:*
3: em_estin:*
4: em_estin:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estl:*
9: gb_estl:*
10: gb_est2:*
11: gb_estc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mus:*
23: em_gss_mam:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	965.8	80.7	1204	11 AK003748	AK003748 Mus muscu
2	954	79.7	954	29 AY408390	AY408390 Homo sapi
3	951.4	79.5	954	29 AY408391	AY408391 Pan trogl
4	949	79.3	1159	11 AK014761	AK014761 Mus muscu

5	828.6	69.2	926	13	BQ934226	BQ934226 AGENCOURT
6	826	69.0	954	29	AY408392	AY408392 Mus muscu
7	708.8	59.2	1000	13	BY704009	BY704009 BY704009
8	692.4	57.8	1011	12	BI456138	BI456138 603172901
9	674.8	56.4	800	14	CF410248	CF410248 CH3#065 G
10	674	56.3	919	14	CB195636	CB195636 AGENCOURT
11	668.4	55.4	961	12	BI694284	BI694284 603347729
12	660.6	54.8	845	14	CB588736	CB588736 AGENCOURT
13	623.8	52.1	803	14	CF410247	CF410247 CH3#065 G
14	623	52.0	932	13	BY714257	BY714257 BY714257
15	620.6	51.8	876	14	CB587595	CB587595 AGENCOURT
16	612.6	51.2	751	14	CD512100	CD512100 AGENCOURT
17	608.2	50.8	867	14	CB587917	CB587917 AGENCOURT
18	579.6	48.4	744	12	BI454468	BI454468 603171428
19	579	48.4	846	10	BF182340	BF182340 601804094
20	577	48.2	958	12	BG174435	BG174435 602334255
21	574	48.0	712	14	CD103859	CD103859 AGENCOURT
22	553.2	46.2	679	12	BI248748	BI248748 602992715
23	531.8	44.4	600	9	AV593197	AV593197 AV593197
24	530.4	44.3	639	14	CD773806	CD773806 AGENCOURT
25	500.6	41.8	735	12	BI694698	BI694698 603347638
26	498.6	41.7	557	12	BI535437	BI535437 398969 MA
27	494	41.3	588	10	AW743952	AW743952 ur24h02.Y
28	493.6	41.2	798	12	BI249999	BI249999 602995241
29	485.8	40.6	610	14	CB440637	CB440637 690932 MA
30	485.4	40.6	636	12	BI304069	BI304069 UI-R-DR0-
31	474.8	39.7	479	9	AI123839	AI123839 qa74e02.x
32	473	39.5	595	13	CB440284	CB440284 690548 MA
33	468	39.1	468	13	BI104255	BI104255 BX104255
34	463.6	38.7	888	10	BF121576	BF121576 601756977
35	460	38.4	462	10	BF439260	BF439260 nab61n07.
36	435.8	36.4	465	14	T12179	T12179 A533F Heart
37	424.6	35.5	856	10	BF123957	BF123957 601759641
38	419	35.0	741	14	CB317710	CB317710 AGENCOURT
39	409.4	34.2	521	13	BX638501	BX638501 BX638501
40	385	32.2	399	9	AI039039	AI039039 ox30b10.s
41	379.8	31.7	1224	14	CB202890	CB202890 AGENCOURT
42	378.6	31.6	733	10	BE375361	BE375361 601230345
43	376.4	31.4	460	9	AI146280	AI146280 qb87b10.x
44	371.6	31.0	468	9	AV593196	AV593196 AV593196
45	369	30.8	433	14	CF112025	CF112025 Shultzomi
46	366.8	30.6	487	9	AI112003	AI112003 UI-R-Y0-m
47	366.2	30.6	504	10	BF523627	BF523627 UI-R-Cl-1
48	360.8	30.1	404	12	BI535434	BI535434 398963 MA
49	358.8	30.0	452	9	AI502787	AI502787 UI-R-Cl-1
50	358.4	29.9	882	10	BF679243	BF679243 602153429
51	352.2	29.4	509	14	CD572281	CD572281 PBL 18 EO
52	350.4	29.3	407	10	BF150118	BF150118 uy82a09.Y
53	336.8	28.1	359	9	AA236166	AA236166 zs42f09.s
54	332.6	27.8	391	10	AW962409	AW962409 EST374482
55	329.6	27.5	419	14	W97621	W97621 mf97e12.r1
56	327.6	27.4	416	13	BY377460	BY377460 BY377460
57	319.2	26.7	434	10	AW144807	AW144807 EST291897
58	312	26.1	452	10	BH831558	BH831558 BH831558
59	309.6	25.9	661	10	BE288641	BE288641 601094042
60	308.4	25.8	337	9	AA297231	AA297231 EST112774
61	307	25.6	424	9	AI600132	AI600132 EST251835
62	306.8	25.6	387	13	BY772480	BY772480 BY772480
63	298.8	25.0	392	13	BY378738	BY378738 BY378738
64	298	24.9	299	9	AI147044	AI147044 gb82404.x
65	297.4	24.8	487	13	BX517933	BX517933 BX517933
66	293.8	24.5	380	10	BF148505	BF148505 uy82a09.x
67	285	23.8	448	13	BY452021	BY452021 BY452021
68	284.8	23.8	860	13	BUI27029	BUI27029 603113730
69	274	22.9	369	10	AW240909	AW240909 uq38a03.x
70	269.8	22.5	542	14	CB114632	CB114632 K-EST0158
71	269.6	22.5	345	13	BY112564	BY112564 BY112564
72	268.8	22.5	349	13	BY778741	BY778741 BY778741
73	268.6	22.4	352	13	BY111994	BY111994 BY111994
74	266.6	22.3	342	13	BY115292	BY115292 BY115292
75	266	22.3	1109	13	BUI131955	BUI131955 603119028
76	263.4	22.0	338	13	BY789211	BY789211 BY789211
77	261.6	21.9	341	13	BY086042	BY086042 BY086042

[illegible]

```
Db      98  CCTCAGAGAACTGTGAGGCTGTCAATCTTAAATGCAGAAAGCTCTGAAATCTAAGAAG 157
Qy      131 ATATGTAATCACTTAAGATTGTGGAATGTGTGTTTGGTATCTCTGCGCCCTAACTCTAATT 190
Db      158 ATATGTAATCACTGAAGATTGTGGAATGTGTGTTTGGTATCTCTGCGCCCTTAACCTTAATT 217
Qy      191 GTCCTGTTTGGGGAGCAAGCACTTCTGGCGGAGGTACCCAAAAAAGCCCTATGACATG 250
Db      218 GTCCTGTTTGGGGAGCAAGCACTTCTGGCGGAGGTATCCAAGAAAAACCTATGACATG 277
Qy      251 GAGCACACTTCTACAGCAATGAGAGAGAGAGATTATCATGGAATTTGATCCTGTG 310
Db      278 GAGCACACTTCTACAGCAATGAGAGAGAGAGATTATCATGGAATTTGATCCTGTG 337
Qy      311 ACCAGAACTGAATATTCAGAAAGCGGAAATGGCACTGATGAAACATTTGGAAGTGCACGAC 370
Db      338 ACCAGAACAGAAATATTCAGAAAGTGGAAATGGCACTGATGAAACATTTGGAAGTGCATGAC 397
Qy      371 TTTAAAAACGATACACTGGCATCTACTTCTGCGGTCTTCAAAAATGTTTATCAAAACT 430
Db      398 TTTAAAAATGGATACACTGGCATCTACTTCTGAGTCTTCAAAAATGCTTTATTAATACT 457
Qy      431 CAGATTAAAGTATTCCTGAATTTTCTGAACCCAGAGAGGAATAGATGCAATGAAGAA 490
Db      458 CAAATCAAGTATTCCTGAATTTTCTGAACCCAGAGGA--AAATAGATGGAATGAAGAA 515
Qy      491 ATTACCAAACTTTCTTTGAACAGTCACTGATTTGGGTCCCGAGAGAAAGCCCTATTGAA 550
Db      516 ATTACTACAACTTTCTTTGAACAGTCACTGATTTGGGTCCCGAGAGAAAGCCCTATTGAA 575
Qy      551 AACCGAGATTTCTTAAAAATTCGAAATTTCTGGAGATTTGTGATAACGTGACCATGTAT 610
Db      576 AACAGAGACTTCTGAAATAATCTTAAATTTCTGGAGATTTGCGATAATGTGACCATGTAT 635
Qy      611 TGGATCAATCCCACTCTAATATAGTTTCTGAACCCAGAGAGGAATAGATGCAATGAAGAA 670
Db      636 TGGATCAATCCCACTCTAATAGTTTCTGAACCCAGAGGAATAGATGCAATGAAGAA 695
Qy      671 GATCTTCACTTTCTCTGCCAACGAAAAAAGGGATTGAAACAAATGAACAGTGGGTGTC 730
Db      696 GATCTTCACTTTCTCTGCCAACGAAAAAAGGGATTGACCAAGATGAGCAATGGGTGTC 755
Qy      731 CCTCAAGTGAAGTAGAGAGACCGCTACGCGAGACAAGCAAGTGAAGGAACCTTCCA 790
Db      756 CCGCAAGTGAAGTGAAGAGACCGCGCACACAGACAAGCAAGCGAGGAACCTTCCCT 815
Qy      791 ATAAATGACTATCTGAAATGGAATAGAAATTTGATCCCATGCTGGATGAGAGAGTTAT 850
Db      816 ATAAATGACTATCTGAAATGGAATTTGAAATTTGACCCAAATGCTGGATGAGAGAGTTAC 875
Qy      851 TGTGTATTTACTGCCGTCGAGGCAACCGCTATTGCGCGCGCTCTGTGAACCTTTACTA 910
Db      876 TGTGTATTTACTGCTGAGGCAACCGTTACTGCGCGCGCTCTGTGNAACCTTTACTA 935
Qy      911 GGTACTACCCATATCCATCTCTACAAAGAGAGCGAGTCAATCTGTGCTGTCATCATG 970
Db      936 GGTACTACCCATACCCCTACTCTACCAAGAGAGTCTGAGTCACTCTGTCTGTCATCATG 995
Qy      971 CTTTGTAACTGGTGGGTGGCCCGCATGCTGGGAGGGTCTTAATAGAGGTTTGAGCTCAA 1030
Db      996 CTTTGTAACTGGTGGGTGGCCCGCATGCTGGGAGAGTCTTAATAGAGAGTTTGAGTCAA 1055
Qy      1031 ATGCTTAAACTGCT----GSCAACATATAATAATGATGCTATTCAATGAATTTCTGCG 1086
Db      1056 ACGCTTAACTTCTGTTAGCCATATATAATTAATGATGCTACTCCATGAATTTCTGCC 1115
Qy      1087 TATGAGGATCTGGGCCCTGGTAGCCAGCTCTCCAGAATTTACTTGTAGGTAATTCCTCTC 1146
Db      1116 TATGAGGATTTGGCTTCCAAAGTAGCCATCTCTTCAGAAATTTACTTGTAGGATATTCCTCTC 1175
Qy      1147 TTCATGCTCTAATAAACTTCTACATATC 1175
Db      1176 TTCATGCTCTAATAAACTTCTACATATC 1204
```

```
RESULT 2
AY408390
LOCUS
DEFINITION
  Homo sapiens HCM3208 gene, VIRTUAL TRANSCRIPT, partial sequence,
  genomic survey sequence.
ACCESSION
  AY408390
VERSION
  AY408390.1 GI:39764361
KEYWORDS
  GSS.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 954)
  Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
  Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
  Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
  Adams,M.D. and Cargill,M.
  Inferring nonneutral evolution from human-chimp-mouse orthologous
  gene trios
JOURNAL
  Science 302 (5652), 1960-1963 (2003)
PUBMED
  14671302
REFERENCE
  2 (bases 1 to 954)
  Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
  Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
  Adams,M.D. and Cargill,M.
  Direct Submission
  Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
  Rockville, MD 20850, USA
  This sequence was made by sequencing genomic exons and ordering
  them based on alignment.
FEATURES
  source
    1..954
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    <1..>954
    /locus_tag="HCM3208"
  gene
  ORIGIN
    Query Match 79.7%; Score 954; DB 29; Length 954;
    Best Local Similarity 100.0%; Pred. NO. 7.9e-173;
    Matches 954; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    Qy 59 ATGCGAAAGAATCTCTCCAGAGAAATTGTGAAGACTGTCAATTTAAATGCAGAAAGCTTTT 118
    Db 1 ATGCGAAAGAATCTCTCCAGAGAAATTGTGAAGACTGTCAATTTAAATGCAGAAAGCTTTT 60
    Qy 119 AATCCAAGAAATATGTAAATCACTTAAGATTGTGGACTGGTGTGTATCCTGGCC 178
    Db 61 AATCCAAGAAATATGTAAATCACTTAAGATTGTGGACTGGTGTGTATCCTGGCC 120
    Qy 179 CTAACTCTAATTGTCTCTGTTTGGGGAGCAAGCACTTCTGGCGGAGGTACCCAAAAA 238
    Db 121 CTAACTCTAATTGTCTCTGTTTGGGGAGCAAGCACTTCTGGCGGAGGTACCCAAAAA 180
    Qy 239 GCCTATGACATGAGCACACTTTCTACAGCAATGGAGAGAAGAAAGATTTACATGGAA 298
    Db 181 GCCTATGACATGAGCACACTTTCTACAGCAATGGAGAGAAGAAAGATTTACATGGAA 240
    Qy 299 ATGATCCTGTGACCAAGAACTGAAATATTTCAGAACGGAAATGGCACTGATGAACATTG 358
    Db 241 ATGATCCTGTGACCAAGAACTGAAATATTTCAGAACGGAAATGGCACTGATGAACATTG 300
    Qy 359 GAAGTGCAGCACTTTAAACCGATACACTGCGCATCTACTCTGGGTCTTCAAAATGT 418
    Db 301 GAAGTGCAGCACTTTAAACCGATACACTGCGCATCTACTCTGGGTCTTCAAAATGT 360
    Qy 419 TTTATCAAACTCAGATTAAAGTGATTTCTGAAATTTCTGAACCCAGAGAGAAATAGAT 478
    Db 361 TTTATCAAACTCAGATTAAAGTGATTTCTGAAATTTCTGAACCCAGAGAGAAATAGAT 420
```

QY 479 GAGATGAGAAATTACACAACTTCTTTGAAAGTCAGTGATTTGGTCCCGACAGAA 538
DB 421 GAGATGAGAAATTACACAACTTCTTTGAAAGTCAGTGATTTGGTCCCGACAGAA 480
QY 539 AAGCTATTGAAACCGAGATTTCTTAAATAATCCAAATTTCTGGAGATTGTGATAAC 598
DB 481 AAGCTATTGAAACCGAGATTTCTTAAATAATCCAAATTTCTGGAGATTGTGATAAC 540
QY 599 GTGACCATGTAATGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAG 658
DB 541 GTGACCATGTAATGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAG 600
QY 659 GAGGGGAGAGAGATCTTCACTTTCTGCCAACGAAAAAGGATTGAACAAATGAA 718
DB 601 GAGGGGAGAGAGATCTTCACTTTCTGCCAACGAAAAAGGATTGAACAAATGAA 660
QY 719 CAGTGGGTGGTCCCTCAAGTGAAGTAGAAGACCCGCTCAGCCAGACAAGCAAGTGAG 778
DB 661 CAGTGGGTGGTCCCTCAAGTGAAGTAGAAGACCCGCTCAGCCAGACAAGCAAGTGAG 720
QY 779 GAAGAACTTCCAAATAATGACTATATCTGAAATGGAATGGAATTTGATCCCATCTGGAT 838
DB 721 GAAGAACTTCCAAATAATGACTATATCTGAAATGGAATGGAATTTGATCCCATCTGGAT 780
QY 839 GAGGAGGTTATTTGTTGTTTACTGCGTTCGAGCAACCGCTATTGCGCGCGCTCTGT 898
DB 781 GAGGAGGTTATTTGTTGTTTACTGCGTTCGAGCAACCGCTATTGCGCGCGCTCTGT 840
QY 899 GAACCTTTACTAGGCTACTACCATATCCATCTGCTACCAAGGAGGAGTCAATCTGT 958
DB 841 GAACCTTTACTAGGCTACTACCATATCCATCTGCTACCAAGGAGGAGTCAATCTGT 900
QY 959 CGTGTCATCATGCTTGTAACTGGTGGTGGCGCGCATGCTGGGAGGGTCTAA 1012
DB 901 CGTGTCATCATGCTTGTAACTGGTGGTGGCGCGCATGCTGGGAGGGTCTAA 954

RESULT 3
AY408391
LOCUS
DEFINITION Pan troglodytes HCM3208 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY408391
VERSION AY408391.1 GI:39764362
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE
AUTHORS
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
Gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..954
/organism="Pan troglodytes"
/mol_type="genomic DNA"

/db_xref="taxon:9598"
<1..>954
/locus_tag="HCM3208"
Query Match 79.5%; Score 951.4; DB 29; Length 954;
Best Local Similarity 99.8%; Pred. No. 2.5e-172; Indels 0; Gaps 0;
Matches 952; Conservative 0; Mismatches 2;
ORIGIN
QY 59 ATGCCAAAGAAATCCTCCAGAGAAATGTGAAGACTGTGCATTTCTAAATGCAAGACTTTT 118
DB 1 ATGCCAAAGAAATCCTCCAGAGAAATGTGAAGACTGTGCATTTCTAAATGCAAGACTTTT 60
QY 119 AAATCCAAAGAAATATATGTAATATCACTTAAGATTTGTGACTGTGTTTGGTATCTGGCC 178
DB 61 AAATCCAAAGAAATATATGTAATATCACTTAAGATTTGTGACTGTGTTTGGTATCTGGCC 120
QY 179 CTAATCTAAATGTCCTGTTTGGGGAGCAAGCACTTCTGSCCGAGGTACCCAAAAA 238
DB 121 CTAATCTAAATGTCCTGTTTGGGGAGCAAGCACTTCTGSCCGAGGTACCCAAAAA 180
QY 239 GCCTATGATGAGGAGCACACACTTTCTACAGCAATGGAGAGAAAGAAATTTACATGAA 298
DB 181 NCCTATGATGAGGAGCACACACTTTCTACAGCAATGGAGAGAAAGAAATTTACATGAA 240
QY 299 ATTGATCTGTGACCAAGAACTGAAATATTCAGAACGGAAATGSCACTGATGAAACATTTG 358
DB 241 ATTGATCTGTGACCAAGAACTGAAATATTCAGAACGGAAATGSCACTGATGAAACATTTG 300
QY 359 GAGTGCAGCACTTTAAATAACGGATACACTGTCATCTTCTGCTGGTCTTCAAAAATGT 418
DB 301 GAGTGCAGCACTTTAAATAACGGATACACTGTCATCTTCTGCTGGTCTTCAAAAATGT 360
QY 419 TTTATCAAAACTCAGATTTAAAGTGTCTCTGAAATTTCTGAAACCAAGAGAGAAATAGAT 478
DB 361 TTTATCAAAACTCAGATTTAAAGTGTCTCTGAAATTTCTGAAACCAAGAGAGAAATAGAT 420
QY 479 GAGATGAGAAATTAACCAAACTTTCTTTGAAACAGTCACTGATTTGGTCCCGACAGAA 538
DB 421 GAGATGAGAAATTAACCAAACTTTCTTTGAAACAGTCACTGATTTGGTCCCGACAGAA 480
QY 539 AAGCTATTGAAACCGAGATTTCTTAAATAATCCAAATTTCTGGAGATTGTGATAAC 598
DB 481 AAGCTATTGAAACCGAGATTTCTTAAATAATCCAAATTTCTGGAGATTGTGATAAC 540
QY 599 GTGACCATGTAATGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAG 658
DB 541 GTGACCATGTAATGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAG 600
QY 659 GAGGGGAGAGAGATCTTCACTTTCTGCCAACGAAAAAGGATTGAACAAATGAA 718
DB 601 GAGGGGAGAGAGATCTTCACTTTCTGCCAACGAAAAAGGATTGAACAAATGAA 660
QY 719 CAGTGGGTGGTCCCTCAAGTGAAGTAGAAGACCCGCTCAGCCAGACAAGCAAGTGAG 778
DB 661 CAGTGGGTGGTCCCTCAAGTGAAGTAGAAGACCCGCTCAGCCAGACAAGCAAGTGAG 720
QY 779 GAAGAACTTCCAAATAATGACTATATCTGAAATGGAATGGAATTTGATCCCATCTGGAT 838
DB 721 GAAGAACTTCCAAATAATGACTATATCTGAAATGGAATGGAATTTGATCCCATCTGGAT 780
QY 839 GAGAGAGTTATTTGTTGTTTACTGCGTTCGAGCAACCGCTATTGCGCGCGCTCTGT 898
DB 781 GAGAGAGTTATTTGTTGTTTACTGCGTTCGAGCAACCGCTATTGCGCGCGCTCTGT 840
QY 899 GAACCTTTACTAGGCTACTACCATATCCATCTGCTACCAAGGAGGAGTCAATCTGT 958
DB 841 GAACCTTTACTAGGCTACTACCATATCCATCTGCTACCAAGGAGGAGTCAATCTGT 900
QY 959 CGTGTCATCATGCTTGTAACTGGTGGTGGCGCGCATGCTGGGAGGGTCTAA 1012
DB 901 CGTGTCATCATGCTTGTAACTGGTGGTGGCGCGCATGCTGGGAGGGTCTAA 954

RESULT 4
AK014761 LOCUS DEFINITION

AK014761 1159 bp mRNA linear HTC 20-SEP-2003
Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched library, Clone:4833424007 product:tenomodulin, full insert sequence.

ACCESSION
VERSION AK014761.1 GI:12852797
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Carninci, P. and Hayashizaki, Y.
JOURNAL High-efficiency full-length cDNA cloning
MEDLINE Meth. Enzymol. 303, 19-44 (1999)
PUBMED 99279253
10349636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to Prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, T., Wataniki, M., Oneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichipillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)

REFERENCE
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1159)

REFERENCE
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sobabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
SUBMITTED Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

JOURNAL TITLE

QY	381	GATACACTGGCATCTACTTCTGTTGCTCTTCAAAATGTTTTATCAAACTCAGATTAAAG	440
Db	361	GATACACTGGCATCTACTTCTGTTGCTCTTCAAAATGTTTTATCAAACTCAGATTAAAG	420
QY	441	TGATTCCTCGAATTTCTGAACAGAGAGGAAATAGATGAGAATGAAGAAATACCAAA	500
Db	421	TGATTCCTCGAATTTCTGAACAGAGGAAATAGATGAGAATGAAGAAATACCAAA	480
QY	501	CTTTCCTTGAACAGTCAGTGATTTGGTCCAGCAGAAAGCCTATTGAAACCGAGATT	560
Db	481	CTTTCCTTGAACAGTCAGTGATTTGGTCCAGCAGAAAGCCTATTGAAACCGAGACT	540
QY	561	TTCTTAAAAATTCCAAAATTCGAGAGATTGFGATAAGCTGACCATGATTTGGATCAATC	620
Db	541	TTCTTAAAAATTCCAAAATTCGAGAGATTGFGATAAGCTGACCATGATTTGGATCAATC	600
QY	621	CCACTCTAATATCAGTTCTGAGTTTCAAGACTTTGAGGAGGAGGAGAGATCTTCACT	680
Db	601	CCACTCTAATAGCAGTTTCAAGACTTTGAGGAGGAGGAGAGATCTTCACT	660
QY	681	TTCTCTGCAACCAAAAAAGGGATTGACAAAAATGAACAGTGGTGGTCCCTCAAGTGA	740
Db	661	TTCTCTGCAACCAAAAAAGGGATTGACAAAAATGAACAGTGGTGGTCCCTCAAGTGA	720
QY	741	AAATGAGAGAACCCGCTCAGCCAGACAGCAAGTGAAGAACTTCCAAATAAATGACT	800
Db	721	AAATGAGAGAACCCGCTCAGCCAGACAGCAAGTGAAGAACTTCCAAATAAATGACT	780
QY	801	ATACGAAATGGAATAGAAATTTGATCCCATGCTGATGAGAGAGGTTATTTGTTATTT	860
Db	781	ATACGAAATGGAATTTGAAATTTGACCAATGCTGATGAGAGAGGTTATTTGTTATTT	840
QY	861	ACTGCGCTGAGGCAACGCTATTGCGCGGCTGTGAACTTTACTAGGCTACTACC	920
Db	841	ACTGCGCTGAGGCAACGCTATTGCGCGGCTGTGAACTTTACTAGGCTACTACC	900
QY	921	CATATCCATCTGCTACCAAGGAGGACGAGTCACTGCTGTCATFCATGCTTGTAACT	980
Db	901	CATACCCCTACTGCTACCAAGGAGGTCGAGTCACTGCTGTCATFCATGCTTGTAACT	960
QY	981	GGTGGTGCCCGCATGCTGGGGAGGGTCTAATAGAGAGGTTGAGCTCAAAATGCTTAAAC	1040
Db	961	GGTGGTGCCCGCATGCTGGGGAGGGTCTAATAGAGAGGTTGAGCTCAAAATGCTTAAAC	1020
QY	1041	TGCT---GGCACATATATAATCATGCTATTCAATGAATTTCTGCTATGAGGCAT	1096
Db	1021	TTCTGTTAGCAATATATAATTAATGCTACTCTCAATGAATTTCTGCTATGAGGCAT	1080
QY	1097	CTGGCCCTGGTAGCCAGCTCTCCAGAAATTAATTTAGGTAATTCCTCTCTTCATGTTCT	1156
Db	1081	TTGCTTCCAAAGTAGCTATCTTCCAGAAATTAATTTAGGTAATTCCTCTCTTCATGTTCT	1140
QY	1157	AATAAATCTTCAATATATC 1175	
Db	1141	AATAAATCTTCAATATATC 1159	
RESULT 5			
BQ934226			
LOCUS			
DEFINITION			
BQ934226 8883116 Lupski sciatic nerve Homo sapiens cDNA clone			
IMAGE:6200216 5', mRNA sequence.			
ACCESSION			
BQ934226			
VERSION			
BQ934226.1 GI:22349609			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 926)			
NIH-MGC http://mgs.nci.nih.gov/			
REFERENCE			
AUTHORS			
TITLE			
National Institutes of Health, Mammalian Gene Collection (MGC)			

JOURNAL COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM13614 row: 1 column: 09
 High quality sequence stop: 640.

FEATURES

source

1..926
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6200216"
 /sex="male"
 /tissue_type="sciatic nerve"
 /dev_stage="adult, 70 yr"
 /lab_host="DH10B"
 /clone_lib="Lupski sciatic nerve"
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1: Not1; Site 2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCAGCGCTCCG-3' and 5'-GACTAGTTCTAGATCGAGCGCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

ORIGIN

Query Match	69.2%;	Score	828.6;	DB 13;	Length	926;
Best Local Similarity	95.7%;	Pred. No.	8.2e-149;			
Matches	864;	Conservative	0;	Mismatches	34;	Indels
						5; Gaps

QY	97	CATTCTAAATGCAGAGCTTTTAAATCCAAGAAAATATGTAATCACTTAAGATTGTGG	156
Db	1	CATTCTAAATGCAGAGCTTTTAAATCCAAGAAAATATGTAATCACTTAAGATTGTGG	60
QY	157	ACTGGTGTGGTATCTGCGCCCTAACTCTAATGTCCTGTTTGGGGAGGACGACCTT	216
Db	61	ACTGGTGTGGTATCTGCGCCCTAACTCTAATGTCCTGTTTGGGGAGGACGACCTT	120
QY	217	CTGGCCGAGGTACCCAAAAGCCCTATGACATGGAGCACACTTCTACAGCAATGGAGA	276
Db	121	CTGGCCGAGGTACCCAAAAGCCCTATGACATGGAGCACACTTCTACAGCAATGGAGA	180
QY	277	GAGAAGAAAGATTACATCGAAATTCATCTGTGACAGAACTGAAATATTCAGAGCGG	336
Db	181	GAGAAGAAAGATTACATCGAAATTCATCTGTGACAGAACTGAAATATTCAGAGCGG	240
QY	337	AAATGCACTGATGAACACTTGAAGTGCACGACTTTTAAAAACGGATACACTGGCATCTA	396
Db	241	AAATGCACTGATGAACACTTGAAGTGCACGACTTTTAAAAACGGATACACTGGCATCTA	300
QY	397	CTTCGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTCCTGAATTTTC	456
Db	301	CTTCGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTCCTGAATTTTC	360
QY	457	TGAACCAAGAGGAAATAGATGAGAAATGAAGAAATTCACCAACTTCTTTGAACAGTC	516
Db	361	TGAACCAAGAGGAAATAGATGAGAAATGAAGAAATTCACCAACTTCTTTGAACAGTC	420
QY	517	AGTGATTTGGTCCCGACAGAAAGCCTATTGAAACCGAGATTCTTAAAAATTCCAA	576
Db	421	AGTGATTTGGTCCCGACAGAAAGCCTATTGAAACCGAGATTCTTAAAAATTCCAA	480
QY	577	AATCTCGAGATTGTGATAACCGTGCACCTGATTTGGATCAATCCCACTCTAATATCAGT	636

Db		481	AATTCTGGAGATTTGTGATAACGTCACCATGPAITGGATCAATCCCACTTAAATATCAGT	544
Qy		637	TTCTGAGTTACAAGACTTTTGAGGAGGAGGGAAGATCTTCACTTTCTGCACACGAAAA	696
Db		541	TTCTGAGTTACAAGACTTTGAGGAGGAGGGAAGATCTTCACTTTCTGCACACGAAAA	600
Qy		697	AAAAGGGATTGACAAAATGAACAGTCGGGTGCTCCCTCAAAGTGAAGAAGTAGAACGCCG	756
Db		601	AAAAGGGATTGACAAAATGAACAGTCGGGTGCTCCCTCAAAGTGAAGAAGTAGAACGCCG	660
Qy		757	TCACGCCACAGCAAGCAAGTGAGGAGAGACTTCCCATTAATGACTATACCTGAAATCGAAT	816
Db		661	TCACGCCACAGCAAGCAAGTGAGGAGAGACTTCCCATTAATGACTATACCTGAAATCGAAT	720
Qy		817	AGAAATTTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTATTTACTGCCGTTCGAGGCAA	876
Db		721	AGAAATTTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTATTTACTGCCGTTCGAGGCAA	780
Qy		877	CGGCTATTTCGCGCGCTCTGTGAACCTTTACTAGGCTACTACCATATTCATATCTG---	933
Db		781	CGGCTATTTCGCGCGCTCTGTGAACCTTTACTAGGCTACTACCATATTCATATCTG---	840
Qy		934	--CTACCAAGGAGGACGAGTCATCTGCTGCTCATCATGCTTGAACCTTGGGAAAACCTTGGGGGGGG	991
Db		841	ACCCAAGGAGGACGAGTCATCTGGGCGGGGCATCATGCTTGGGAAAACCTTGGGGGGGG	900
Qy		992	CGC 994	
Db		901	GGC 903	
RESULT 6				
AY408392				
LOCUS				
DEFINITION			954 bp DNA linear GSS 12-DEC-2003	
ACCESSION			Mus musculus HCM3208 gene, VIRTUAL TRANSCRIPT, partial sequence,	
VERSION			AY408392	
KEYWORDS			AY408392.1 GI:39764363	
SOURCE			GSS.	
ORGANISM			Mus musculus (house mouse)	
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS			Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
			1 (bases 1 to 954)	
			Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,	
			Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,	
			Ferriera,S., Wang,G.; Zheng,X.H., White,T.J., Sninsky,J.J.,	
			Adams,M.D. and Cargill,W.	
			Inferring nonneutral evolution from human-chimp-mouse orthologous	
			gene trios	
JOURNAL			Science 302 (5652), 1960-1963 (2003)	
PUBMED			14671302	
REFERENCE			2 (bases 1 to 954)	
AUTHORS			Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,	
			Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,	
			Ferriera,S., Wang,G.; Zheng,X.H., White,T.J., Sninsky,J.J.,	
			Adams,M.D. and Cargill,M.	
			Direct Submission	
TITLE			Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,	
JOURNAL			Rockville, MD 20850, USA	
COMMENT			This sequence was made by sequencing genomic exons and ordering	
			them based on alignment.	
FEATURES			Location/Qualifiers	
source			1..954	
			/organism="Mus musculus"	
			/mol_type="genomic DNA"	
			/db_xref="taxon:10090"	
gene			<1..954	
			/locus_tag="HCM3208"	
ORIGIN				
Query Match			69.0%; Score 826; DB 29; Length 954;	
Best Local Similarity			91.6%; Pred. No. 2.6e-148;	

VERSION	BY704009.1	GI:27115128	
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 1000)		
	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Giesi, C., Godzik, A., Gough, J., Jarvis, E.D., Kanai, A., Gustingich, S., Hirokawa, N., Jackson, I.J., King, B.L., Konagaya, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalon, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Shirakawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, M., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
MEDLINE	22354683		
PUBMED	1246851		
COMMENT	<p>Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@sc.riken.go.jp/ URL: http://genome.gsc.riken.go.jp/ Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (http://genome.gsc.riken.go.jp) for</p>		

FEATURES	Location/Qualifiers	source
	1..1000	
	/organism="Mus musculus"	
	/mol_type="mRNA"	
	/strains="C57BL/6J"	
	/db_xref="taxon:10090"	
	/clone="1110017101"	
	/tissue_type="whole body"	
	/dev_stage="18-day embryo"	
	/clone_lib="RIKEN full-length enriched, 18-day embryo whole body"	
ORIGIN		
	Query Match 59.2%; Score 708.8; DB 13; Length 1000;	
	Best Local Similarity 88.1%; Pred. No. 7e-126;	
	Matches 849; Conservative 0; Mismatches 103; Indels 12; Gaps 7;	
QY	11 CTCTAGTCCTCTCAAGCAAGGAAGAGTACTGTGCTGAGAGACCATGGCAAGAAT 70	
Db	38 CTCTAGTCCTCTCAAGCAGGGAAGAGCACGCTGCTGGAGACCATGGCAAGAAT 97	
QY	71 CTCTCAGAGAAATGTGAAGACTGTGCATTTCTAATTCAGAAAGCTTTTAAATCCAAGAA 130	
Db	98 CTCTCAGAGAACTGTGAGGCTGTGCATTTCTAATTCAGAAAGCTTTGAAATCTAAGAAG 157	
QY	131 ATATGTAATTCATTAAGATTTGTGACTGTGTGTTGGTATCTCTGGCCCTTAATCTTAAT 190	
Db	158 ATATGTAATTCATTAAGATTTGTGACTGTGTGTTGGTATCTCTGGCCCTTAATCTTAAT 217	
QY	191 GTCTCTGTTTGGGGAGCAGCACTTCTGCGCGAGGTACCCAAAAAAGCCTATGACATG 250	
Db	218 GTCTCTGTTTGGGGAGCAACACTTCTGCGCGAGGTATCCAAAGAAACCTATGACATG 277	
QY	251 GAGCACACTTTCTACAGCAATGGAGAGAAGAAGATTTACATGAGAAATGATCTCTGTG 310	
Db	278 GAGCACACTTTCTACAGCAACGCGAGAGAAGAAGATTTACATGAGAAATGATCTCTG 337	
QY	311 ACCAGAACTGAAATATTTACAGAACGGAATGGCACTGATGAAACATTTGGAAGTCAGAC 370	
Db	338 ACCAGAACTGAAATATTTACAGAACGGAATGGCACTGATGAAACATTTGGAAGTCAGAC 397	
QY	371 TTTAAAAACGGATACACTGGCATCTACTCTGCGGTCTTCAAAAATGTTTATCAAAACT 430	
Db	398 TTTAAAAACGGATACACTGGCATCTACTCTGCGGTCTTCAAAAATGTTTATCAAAACT 457	
QY	431 CAGATTTAAAGTGATTCCTGAATTTTCTGAACCGAGAGGAATAGATGAGAAATGAAGAA 490	
Db	458 CAAATCAAGTGATTCCTGAATTTTCTGAACCGAGAGGAATAGATGAGAAATGAAGAA 515	
QY	491 ATTACCAAACTTTCTTTGAACAGTCACTGATTTGGTCCAGCAGAAAAAGCCTATTGAA 550	
Db	516 ATTACTACAACTTTCTTTGAACAGTCACTGATTTGGGTCCCGAGAAAAAGCCTATTGAA 575	
QY	551 AACCGAGATTTCTTAAAAATTTCAAAAATTTCTGAGATTTGTGATAAGTCAGCATGTAT 610	
Db	576 AACGAGACTTCCTGAAAAATTTCTAAAAATTTCTGAGATTTGTGATAAGTCAGCATGTAT 635	
QY	611 TGGATCAATCCCACTTAATATCACTTTCTGAGTTTACAAGCTTTGAGGAGGAGGAGAA 670	
Db	636 TGGATCAATCCCACTTAATATCACTTTCTGAGTTTACAAGCTTTGAGGAGGAGGAGAA 695	
QY	671 GATCTTCACTTTCTGCGCAAGAAAAAGGGATTTGAACAAAAATGAACAGTGGGTGGTC 730	
Db	696 GATCTTCACTTTCTGCGCAAGAAAAAGGGATTTGAACAAAAATGAACAGTGGGTGGTC 755	
QY	731 CCTCAAGTGAAGTAGAGAGACCGCTCACGCCAGACAAAGCAAGTGAAGAAATCTTCCA 790	
Db	756 CCGCAAGTGAAGTAGAGAGACCGCTCACGCCAGACAAAGCAAGTGAAGAAATCTTCCA 814	
QY	791 ATAAATGACTACTGAAAAATGGAATAGAAATTTGATCCCATGCTGGATGAGAGAGTTAT 850	
Db	815 ATAACTGACTACTGAAAAATGGAATAGAAATTTGATCCCATGCTGGATGAGAGAGTTAT 874	

```

QY 851 TGTGTGATTTACTGCGTGCAGCAACCGCTATTGCGCGGCTGTGTGAACCTTTACTA 910
D 875 --GTGTTTACTGTGCTGCAG--AAACGGTACTGCGCGG-GTCTGTGNACNTTTACTA 929
QY 911 GGTACTACTCCATATCCATACCTACTGCTACCAAGAGAGACGAGTCACTGTGTCATCATG 970
D 930 GCTATACACATA---CCCTACTGCTACAAAGGNGGTGAGTC-TCTGTGTCGTATCATG 985
QY 971 CCTT 974
D 986 CTT 989

BI456138 1011 bp mRNA linear EST 21-AUG-2001
LOCUS 603172901F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5252233 5',
DEFINITION mRNA sequence.
ACCESSION BI456138
VERSION BI456138.1 GI:15246794
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11637 row: a column: 02
High quality sequence stop: 810.
FEATURES
    source
        1..1011
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /clone="IMAGE:5252233"
            /tissue_type="tumor, gross tissue"
            /dev_stage="7 months"
            /lab_host="DH10B"
            /clone_lib="NCI CGAP Mam5"
            /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sall;
            Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
            Library constructed by Life Technologies. Investigators
            providing samples: Lothar Hennighausen/Robin Humphreys,
            NIH"
    ORIGIN
        Query Match 57.8%; Score 692.4; DB 12; Length 1011;
        Best Local Similarity 85.3%; Pred. No. 9.6e-123;
        Matches 856; Conservative 0; Mismatches 136; Indels 12; Gaps 7;

QY 32 GGAAGAGTACTGTGTGCTGAGAGACCATGCAAGAATCTCCAGAGAATTGTGAAGAC 91
D 1 GGAAGAGACCGTGTGCTGGAGACCATGCAAGAATCTCCAGAGAATTGTGAGGGC 60
QY 92 TGTCAATTTAAATGACAGAGCTTTTAAATCCAGAAAATATGTAATCACTTAAGATT 151
D 61 TGTCAATTTAAATGACAGAGCTCTGAATCTAAGAAGATATGTAATCACTGAAGATT 120
QY 152 TGTGGCTGCTGTTGTATCTCTGGCCCTAACTCAATTCCTGTTTGGGGAGCAAG 211

```

```

D 121 TGTGACTAGTGTGTTGGTATCTGCTCTAACTTAATTTGTCTTG-TTTGGGGAGCABA 179
QY 212 CACTTCTGCGCGAGGTACCCAAAAAGCCTATGATGATGGAGACACACTTTCTACAGCAAT 271
D 180 CACTTCTGCGCGAGGTATCCAAAGAAAACCTATGATGGAGACACACTTTCTACAGCAAC 239
QY 272 GGAGAGAAGAAGAGATTTCATATGAAAATTGATCTGTGTGACCAGAACTGAAATATTGACA 331
D 240 GCGAGAGAGAGAGAGATTTCATGGAATTTGATCCATTAACCCAGAACAGAAATATTGACA 299
QY 332 AGCGGAAATGGCACTGATGAAAACATTGGAAGTGCACGACTTTTAAAAACGGATACACTGGC 391
D 300 AGTGAATATGGCACTGATGAAAACATTGGAAGTCCATGACTTTTAAAAATGGATACACTGGC 359
QY 392 ATCTACTTCTGCGGTCTTTCAAAAATGTTTATCAAAAACCTAGATTAAAGTGAATTCCTGAA 451
D 360 ATCTACTTCTGAGGTCTTTCAAAAATGCTTTTATTAATACTCAAAATCAAAAGTGAATTCCTGAA 419
QY 452 TTTTCTGAACCCAGAGAGAGAAATAGATCAGAAATGAAGAAATTTACCCACAACTTTCTTTGAA 511
D 420 TTTTCTGAACCCAGAGAGAGAAATAGATCAGAAATGAAGAAATTTACTACAACTTTCTTTGAA 479
QY 512 CAGTCAGTATTGCGTCCCGACAGAAAAGCCTATTGAAAACCCGAGATTTTCTTAAAAAT 571
D 480 CAGTCAGTATTGCGTTCGCCGACAGAAAAGCCTA-TGAAAACAGAGACTTCTCTGAAAAAT 538
QY 572 TCCAAATTTCTGGAGATTGTGTAAACGTGACCATGTA-TTGGATCAATCCCACTCTTAAT 630
D 539 TCTAAAATTTCTGGAGATTGCGATAATGTGACCATGTACTCGGATCAATCCCACTCTTAAT 598
QY 631 ATCAGTCTTCTGAGTTTACAAGACTTTTGAGGAGGAGGAGAGATCTTCACTTTCTTGCCTAA 690
D 599 AGCAGTTTCAGAAATTA CAGACTTTTGAGGAGGAGCGTGAAGATCTTCACTTTCTTACCAG 658
QY 691 CGAAAAAAAGGGATTGAACAAAATAACACAGTGGGTGCTCCCTCAAGTGAAGTAGAGAA 750
D 659 TGAAAAAAAGGGGATTGACCAAGATGAGCAATGGTGT-CCGCAAGTGAAGTGGAGAA 717
QY 751 GACCGGTCAAGCCAGACAGCAAGTGAAGAGAACTTCCAAATAAATGACTATATCTGAAAA 810
D 718 GACC---GCAACACAGCAAGCAAGCAGGAGAGAACTTCTCTATAAATGACTATATCTGAAAA 774
QY 811 TGAATAGAAATTTGATCCCATGCTGGATGAGAGAGTTATTTGTTATTTACTGCGCTCG 870
D 775 TGGAAATGAAATTTGACAAATGCTGGATGAGAGAGGTTACTCTGTGTATTTACTGTGAGTC 834
QY 871 AGGCAACCGCTATT---GCCGCCGCTGTGTGAACCTTT-CTAGGCTACTACCCATAT 925
D 835 GAGGGCAACCGGTTACTGCGCGGCTGTCTGTGAACCTTTAACTAGGTATTAAACCAATTA 894
QY 926 CCATCTCTCTACCAAGGAGGAGACGAGTCACTGTGTGTCATCATGCTTGTAACTGGTGG 985
D 895 CCCTAATGGTACCAAGGAGTCCAGGTCCTCTGCGGAGGACTCCATGCTTGGCAGCTGGTGG 954
QY 986 GTGCGCCCGCATGCTGGGAGGCTTAATAGAGGTTTGAGCTCA 1029
D 955 GTGCGCCGATTTGCTGGGAGAGTTCTTAAAGGAAGTGAAGTTTCA 998

```

```

RESULT 9
CF410248
LOCUS
DEFINITION
    CH3#065_G0177 Canine heart nonnormalized cDNA Library in pBluescript
    Canis familiaris cDNA clone CH3#065_G01 5', mRNA sequence.
ACCESSION
    CF410248
VERSION
    CF410248.1 GI:34411494
KEYWORDS
    EST.
SOURCE
    Canis familiaris (dog)
    Canis familiaris
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
    1 (bases 1 to 800)
    Yi, Y., Desai, R., Olarte, M., Henthorn, P. and George A.L.

```

TITLE Expressed sequence tags from Canine heart
JOURNAL Unpublished (2003)
COMMENT Other_ESTs: CH3#065_G01T3

Contact: George AL

Division of Genetic Medicine

Vanderbilt University

529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232-0275, USA

Tel: 615 936 2660

Fax: 615 936 2661

Email: al.george@vanderbilt.edu

Insert length: 1647 Std Error: 0.00

Seq primer: T7: TAATACGACTCACTATAGGG

High quality sequence start: 39

High quality sequence stop: 797.

Location/Qualifiers

1. .800

/organism="Canis familiaris"

/mol_type="mRNA"

/db_xref="taxon:9615"

/clone="CH3#065_G01"

/tissue_type="heart"

/cell_type="heart"

/dev_stage="mixed developmental stages (adult, 30 day - 40

day fetal)"

/clone_lib="Canine heart normalized cDNA Library in

pBluescript"

/note="Organ: heart; Vector: pBluescript; Site: 1: 5' of

vector NotI; Site: 2: 3' of vector EcoRI; tissue source:

dog heart (adult, 30 day - 40 day fetal), right and left

atria and ventricle. Dog breed - mixed (beagle, German

shepherd, pointer, Irish setter). Library construction:

oligo-dT primed"

ORIGIN

Query Match 56.4%; Score 674.8; DB 14; Length 800;

Best Local Similarity 93.7%; Pred. No. 2.4e-119;

Matches 714; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY 42 CTGTGTCTGAGAGACCATGGCAAGAACTCTCCAGAGAAATGTGAAGACTGTCAATTC 101

DB 39 CTGTGTCTGAGAGACATGGCAAGAACTCTCCAGAGAACTGTGAGGACTGTCAATTC 98

QY 102 TAAATGCAGAAGCTTTTAAATCCAGAGAAATATGTAATCACTTAAGATTGTGGACTGG 161

DB 99 TAAATGCAGAAGCTTTTAAATCCAGAGAGATATGTAATCACTTAAGATTGTGGATTGG 158

QY 162 TGTGTGTATCTCGGCCCTAACTCTAATGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 221

DB 159 TGTGTGTATCTCGGCCCTAACTCTAATGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 218

QY 222 CGGAGGTACCAAAAAGCCTATGACATGAGACACATTTCTACAGCAATGGAGAGAAGA 281

DB 219 CGGAGACACCAAAAAGCAATATGACATGAGACACATTTCTACAGCAATGGAGAGAAGA 278

QY 282 AGAAGATTATCATGGAAATTCATCTCTGTGACCAAGACTGAATAATTCAGAAAGCGGAATG 341

DB 279 AGAAGATTATCATGGAAATTCATCTCTGTGACCAAGACTGAATAATTCAGAAAGTGGAAATG 338

QY 342 GCACATGATGAACATTTGAAGTGCAGACATTTAAACCGGATACACATGGCATCTCTTCG 401

DB 339 GCACATGATGAACATTTGAAGTGCAGATTTTAAATAATGGTTACATGGCATCTCTTCG 398

QY 402 TGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAGTGAATTCCTGAAATTTCTGAAC 461

DB 399 TAGTCTTCAAAAATGCTTCATCAAACTCAGATTAAGTGAATTCCTGAAATTTCTGAAC 458

QY 462 CAGAAGAGGAAATAGATGAGAAATGAGAAATTAACACAACTTCTTGAACAGTCAGTGA 521

DB 459 CAGAAGAGGAAATAGATGAGAAATGAGAAATTAACACAACTTCTTGAACAGTCAGTGA 518

QY 522 TTTGGTCCCGACGAGAAAGCCTATTGAAACCGGATTTTCTTAAATAATTCGAAATTC 581

DB 519 TTTGGTCCCGACGAGAAAGCCTATTGAAATTCGAGACTTTCTGAAATAATTCGAAATTC 578

QY 582 TGGAGATTGTGATTAACCTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCG 641

DB 579 TGGAAATTTGTGATTAATGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCG 638

QY 642 AGTTTACAGACTTTTGGAGGAGGAGAGATCTTTCACTTTCTGCAACGAAAAAAG 701

DB 639 AGTTTACAGACTTTTGGAGGAGGAGAGATCTTTCACTTTCTGCAACGAAAAAAG 698

QY 702 GGATTGAACAAAATGAACAGT-GGGTGGTCCCTCAAGTGAAGTGAAGAGACCCGTCAC 760

DB 699 GTATTGAACAAAACGACGAGTGGGTGGTCCCAAGTGAAGATGAAAAAGATCCGCCAC 758

QY 761 GCCAGACAAAGCAAGTGAAGAACTTCCAAATAATGACTAT 802

DB 759 ACCAGACAAAGCAAGTGAAGAACTTCCAAATAATGACTAT 800

RESULT 10

CB195636

LOCUS

DEFINITION

IMAGE:30137293 5', mRNA sequence.

ACCESSION

CB195636

VERSION

CB195636.1

KEYWORDS

EST.

SOURCE

ORGANISM

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 919)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cga@bbs-r@mail.nih.gov

Tissue Procurement: Dr. David Rowe

cDNA Library Preparation: Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: NDAM0038 row: e column: 14

High quality sequence stop: 662.

Location/Qualifiers

1. .919

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:30137293"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 135"

/notes="Vector: pCMVSPORT6.1; Site 1: EcoRV; Site 2: NotI;

Normalized full-length enriched library from pooled mouse

embryonic limb, maxilla and mandible, day 12.5, 13.5,

14.5, and 15.5 (size selected for the 0.5-1 kb fragments)

Cloned directionally, priming method: Oligo-dT. cDNA

enrichment: >1k bp, Average insert size 1.6k bp.

Normalization (Cot value): 7.5 kb. Priming sequence:

5'-GACATGTTCTAGATCGGAGCGGCCGCTTT-3' Tissue contributed

by, David Rowe. Library constructed by ResGen, Invitrogen

Corp."

ORIGIN

Query Match

Best Local Similarity 56.3%; Score 674; DB 14; Length 919;

Matches 748; Conservative 0; Mismatches 86; Indels 3; Gaps 2;

QY 11 CTCTCAGTCTCTCAAAAGCAAGGAAAGATCTGTGTCTGAGAGACCATGCAAGAAAT 70

DB 13 CTCTCAGTCTCTCAAAAGCAAGGAAAGACCGCTGTGTCTGAGAGACCATGCAAGAAAT 72

```
QY 71 CCTCCAGAGAAATTGTGAAGACTGTCATCTCTAAATGCGAAGAGCTTTTAAATCCAAAGAAA 130
Db 73 CTTCCAGAGAACTGTGAGGCTGTCAATCTAAATGCGAAGAGCTCTGAAATCTAAGAAG 132
QY 131 ATATGTAATACCTTAAGATTTGTGACACTGGTGTGTTGGTATCTGCGCCCTTAACCTCTAATT 190
Db 133 ATATGTAATACCTGAAGATTTGTGACACTAGTGTGTTGGTATCTGCGCCCTTAACCTCTAATT 192
QY 191 GTCTCTTTTGGGGAGCAGCACTCTGCGCGGAGGTACCCAAAAAAGCCCTATGACATG 250
Db 193 GTCTCTTTTGGGGAGCAGCACTCTGCGCGGAGGTATCCAGAAACCTCTATGACATG 252
QY 251 GAGCACATTTCTACAGCAATCGAGAGAGAGAGATTTACATGAAATGATCCTGTG 310
Db 253 GAGCACATTTCTACAGCAACGCGGAGAGAGAGATTTACATGAAATGATCCTCAT 312
QY 311 ACCAGAACTGAAATATTACAGACGCGAATGCGACCTGATGAACATTTGAAAGTGCACGAC 370
Db 313 ACCAGAACGAATATTGAAAGTGAATGCGACTGATGAACATTTGAAAGTGCATGAC 372
QY 371 TTTAAAAACGGATACACTGGCATCTACTTCTGCGGTCTTTCAAAAAATGTTTATCAAAACT 430
Db 373 TTTAAAAATGGATACACTGGCATCTACTTCTGAGTCTTCAAAATGCTTTTATTAAGT 432
QY 431 CAGATTAAAGTGAATCTCTGAATTTCTGAACAGAGAGAAATAGATGAGAAATGAAGAA 490
Db 433 CAATCAAGTGAATCTCTGAATTTCTGAACAGAGAGAAATAGATGAGAAATGAAGAA 492
QY 491 ATTACCAACATTTCTTTGAACAGTCACTGATTTGCGGTCTTTCAAAAAATGTTTATCAAACT 550
Db 493 ATTACTCAACATTTCTTTGAACAGTCACTGATTTGCGGTCTTTGCGGTCTTTGCGGTCTTTGAA 552
QY 551 AACCGAGATTTCTTTAAAAATTTCAAAATTTCTGAGATTTGTGATACTGACCATGAT 610
Db 553 AACAGAGATTTCTTTAAAAATTTCTGAATTTCTGGAGATTTGGGATATGACCATGAT 612
QY 611 TGGATCAATCCCATCTAATATCACTTTCTGAGTTAACAAGTCTTTGAGGAGAGAGAGAA 670
Db 613 TGGATCAATCCCATCTAATATGACAGTTTTCAGAAATTTACAGGACTTTGAGGAGAGAGAGAA 672
QY 671 GATCTTCACTTTCTGCGCAACGAAAAAAGGGATTTGAACAAATGAAACAGTGGGTGTC 730
Db 673 GATCTTCACTTTCTGCGCAACGAAAAAAGGGATTTGAACAAATGAAACAGTGGGTGTC 732
QY 731 CCTCAAGTGAAGTAG-AGAAGACCCGTCAAGCAGACCAAGCAAGTGAAGAGAACTTCC 789
Db 733 CCGCAAGTGAAGTAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 792
QY 790 AATAAATGACTAATGAAATG--AATGAAATTTGATCCATGCTGGATGAGAGA 844
Db 793 CTATATGACTATCTACTGAAATGGAATTTGAATTTTGAACCCATGCTTGAATGAA 849

RESULT 11
LOCUS BI694284
DEFINITION 60347729F1 NCI_CGAP_Mam2 Mus musculus cdna clone IMAGE:5375335 5',
mRNA sequence.
ACCESSION BI694284
VERSION BI694284.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
```

```
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11955 row: j column: 08
High quality sequence stop: 892.
FEATURES
    source
        1..961
            location/Qualifiers
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="FVB/N-3"
                /db_xref="taxon:10090"
                /clone="IMAGE:5375335"
                /tissue_type="tumor, biopsy sample"
                /dev_stage="5 months"
                /lab_host="DH10B"
                /clone_lib="NCI CGAP Mam2"
                /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                Library constructed by Life Technologies. Investigator
                providing samples: Gilbert Smith, NIH"
```

ORIGIN

```
Query Match 55.8%; Score 668.4; DB 12; Length 961;
Best Local Similarity 89.5%; Pred. No. 3.9e-118;
Matches 764; Conservative 0; Mismatches 86; Indels 4; Gaps 4;
QY 240 CCTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAGAAAGATTTACATGGAAA 299
Db 41 CCTATGACATGGAGCACACTTTCTACAGCAACGGGAGAGAGAAAGATTTACATGGAAA 100
QY 300 TTGATCTGTGACGAGAACTGAAATATTTCAGAGCGGAAATGGCACTGATGAACATTTG 359
Db 101 TTGATCCCAATACCAAGACAGAAATATTTCAGAGTGGAAATGGCACTGATGAACATTTG 160
QY 360 AAGTGCAGCACTTTAAACCGGATACACTGCGCATCTACTTCTGTTGGTCTTCAAAATGTT 419
Db 161 AAGTCCATGACTTTAAATGAATACACTGCGCATCTACTTCTGTTAGTCTTCAAAATGCT 220
QY 420 TTATCAAACTCAGATTAAGTGAATTTCTGTAATTTCTGACCAAGAGAGAGAAATAGATG 479
Db 221 TTATTAACCTCAATCAAGTGATTTCTGTAATTTCTGACCAAGAGAGAGAAATAGATG 280
QY 480 AGAATGAAGAAATTAACCAACTTTCTTTGAACAGTCACTGTTGGTCCCGAGCAGAAA 539
Db 281 AGATGAAGAAATTAACCAACTTTCTTTGAACAGTCACTGTTGGTCCCGAGCAGAAA 340
QY 540 AGCCTATTGAAACCGGAGATTTCTTAAATTTCCAAATTTCTGGAGATTTGTGATACG 599
Db 341 AGCCTATTGAAACCGGAGATTTCTTAAATTTCCAAATTTCTGGAGATTTGTGATACG 400
QY 600 TGACCATGTTATGGATCAATCCCACTCTAATATCAGTTTCTGAGTTTACAAGACTTTTCAGG 659
Db 401 TGACCATGTTATGGATCAATCCCACTCTAATATCAGTTTCTGAGTTTACAAGACTTTTCAGG 460
QY 660 AGAGGAGAGAGATCTTCTCTTCCGCAACGAAAGAAAGGGATTTGAACAAAATGAAC 719
Db 461 AGGACGTTGAAGATCTTCTCTTCCACAGTCAAAAGAAAGGGATTTGACCAAGATGAGC 520
QY 720 AGTGGGTGGTCCCTCAAGTGAAGTAGAGNAGACCCGTACCGCAGACAAAGAGTGAAG 779
Db 521 AATGGGTGGTCCCGCAAGTGAAGTAGAGNAGACCCGTACCGCAGACAAAGAGTGAAG 580
QY 780 AAGAACTTCCAAATAAATGACTATCTAGAAATGGAATAGAAATTTGATCCCATCTGATG 839
Db 581 AAGACCTTCTATATATGACTATCTAGAAATGGAATAGAAATTTGATCCCATCTGATG 640
QY 840 AGAGAGTTATTTGTTATTTACTTCCGTCGAGGCAACCGCTATTTCGCCCGCGCTCTGTG 899
Db 641 AGAGAGTTTACTTGTGTTATTTACTTGTGCGTGGAGCAACCGTTACTTCCCGCGCTCTGTG 700
QY 900 AACCTTTACTAGGCTACTACCCATATCCATCTGCTACCAAGAGAGAGAGTCAATCTGTC 959
```

Db	701	AACTTTTACTAGGCTTACCTACCCCTACTGATACAGGAGGTGAGTCACTGTC	760
Qy	960	GTGTCATCATGCTTTGTAACCTGGT-GGGTGGCCCGCATGC-TGGGGAGGGTCTAATAGG-	1016
Db	761	GTGTCATCATGCTTTGCAACTGGTGGGGTGGCCCGCATGCTTGGGAAGAGTCTAATAGGC	820
Qy	1017	AGTTTGTAGCTCAATGCTTAACTGCTGGCA-ACATATAAATGAATGCTGCTATTCAAT	1075
Db	821	AAAGTTGAGTTCAAAACGCTTAACTCTTGTGGTGGTATTTATTATGATGCTGCTACTCCCT	880
Qy	1076	GAATTTCTGCTTAT 1089	
Db	881	GGATTTCTGCTTAT 894	
RESULT 12			
LOCUS	CB588736	845 bp mRNA linear EST 03-APR-2003	
DEFINITION	AGENCOURT 12567644 NIH_MGC_136 Mus musculus cDNA clone		
ACCESSION	CB588736		
VERSION	CB588736.1	GI:29506592	
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 845)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: Dr. David Rowe		
	cDNA Library Preparation: Invitrogen Corp		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: NDAM331 row: j column: 04		
	High quality sequence stop: 738.		
FEATURES	Location/Qualifiers		
source	1..845		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:30290619"		
	/tissue_type="embryonic limb, maxilla and mandible"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_lib="NIH_MGC_136"		
	/note="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;		
	Normalized, full-length enriched library from pool of		
	mouse embryonic limb, maxilla and mandible, embryonic day		
	17.5, 18.5 and newborn (mandible (5, 4 and 1 limb and jaw		
	equivalents from respective days). Cloned directionally,		
	oligo-dr primed (5'-GACTAGTCTAGTCGGAGCGGCCCTT)15-3'.		
	Size selected for the >1kb fragments, average insert size		
	1.2 kb. Normalization to Cot 7.5. Tissue contributed by		
	David Rowe; library constructed by ResGen, Invitrogen		
	Corp. Note: this is a NIH_MGC Library."		
ORIGIN			
Query Match	54.4%; Score 650.6; DB 14; Length 845;		
Best Local Similarity	89.0%; Pred. No. 1e-114;		
Matches 712; Conservative	0; Mismatches 84; Indels 4; Gaps 1;		
Qy	402	TGGGTCTTCAAAAATGTTTATCAAACTCAGATTAAAGTGATCTCTGAATTTTCTGAAC	461
Db	1	TAGGTCTTCAAAAATGCTTTATTAAACCTCAATCAAAAGTGATCTCTGAATTTCTGAAC	60
Qy	462	CAGAGAGGAAATAGATGAGATGAAGAATTAACCAACTTTCTTTGAACAGTCAGTGA	521

Db	61	CAGAGGAAAGAAATAGATGAGAAATGAAGAAATTACTACAACTTTCTTTGAACAGTCAGTGA	120
Qy	522	TTTGGTCCCGCAGCAGAAAAGCCTATTGAAAAACCGAGATTTTCTTAAAAATTCCAAATTC	581
Db	121	TTTGGTTCCTCCGACAGAAAAGCCTATTGAAAACAGAGACTTCTCTGAAAAATTTCTAAAATTC	180
Qy	582	TGGAGATTGTGATACCTGACCATGTTATTTGATCAATCCCACTCTAATATCAGTTTCTG	641
Db	181	TGGAGATTGGGATTAATGTGACCATGTACTGGATCAATCCCACTCTAATATCAGTTTCTAG	240
Qy	642	AGTTACAGACTTTGAGGAGGAGGAGAGATCTTCACCTTTCTGCAACGAAAAAAG	701
Db	241	AATTACAGACTTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	300
Qy	702	GGATTGAACAAAAATGAACAGTGGGTGGTCCCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAG	761
Db	301	GGATTGAACAAAAATGAACAGTGGGTGGTCCCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAG	360
Qy	762	CCAGACAAGCAAGTGAAGGAGAACTTCCAAATAAATGACTATCTGAAATGGAATGAAT	821
Db	361	CCAGACAAGCAAGTGAAGGAGAACTTCCAAATAAATGACTATCTGAAATGGAATGAAT	420
Qy	822	TTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTTTACTTGCCTCGAGGCAACCGCT	881
Db	421	TTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTTTACTTGCCTCGAGGCAACCGCT	480
Qy	882	ATTGCGCGCGCTCTGTAACCTTTACTAGGCTTACTACCCATATCCATATCTGCTACAAG	941
Db	481	ACTGCGCGCGCTCTGTAACCTTTACTAGGCTTACTACCCATATCCATATCTGCTACAAG	540
Qy	942	GAGGACAGTCACTCTGCTGTCATCATGCTTGTAACTGCTGCTGCTGCTGCTGCTGCTGCTG	1001
Db	541	GAGGACAGTCACTCTGCTGTCATCATGCTTGTAACTGCTGCTGCTGCTGCTGCTGCTGCTG	600
Qy	1002	GGAGGCTCTAATAGGAGGTTTGAGCTCAATGCTTAACTGCTTAACTGCTTAACTGCTTAACT	1057
Db	601	GGAGGCTCTAATAGGAGGTTTGAGCTCAATGCTTAACTGCTTAACTGCTTAACTGCTTAACT	660
Qy	1058	AAATGATGCTATTCAATGAATTTCTGCTTATGAGGATCTGAGGATCTGAGGATCTGAGGATCT	1117
Db	661	TAATGATGCTATTCAATGAATTTCTGCTTATGAGGATCTGAGGATCTGAGGATCTGAGGATCT	720
Qy	1118	TCCAGAAATTAATGATGATTAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1177
Db	721	TTCAGAAATTAATGATGATTAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	780
Qy	1178	CAAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA	1197
Db	781	CAAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA	800
RESULT 13			
LOCUS	CF410247	803 bp mRNA linear EST 02-SEP-2003	
DEFINITION	CH3#065_G0173 Canine heart normalized cDNA Library in phaluescript		
ACCESSION	CF410247		
VERSION	CF410247.1	GI:34411493	
KEYWORDS	EST.		
SOURCE	Canis familiaris (dog)		
ORGANISM	Canis familiaris		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
TITLE	1 (bases 1 to 803)		
JOURNAL	Xi.Y., Desai, R., Olarte, M., Henthorn, P. and George A.L.		
COMMENT	Expressed sequence tags from Canine heart		
	Unpublished (2003)		
	Other ESTs: CH3#065_G0177		
	Contact: George AL		
	Division of Genetic Medicine		
	Vanderbilt University		
	529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232-0275, USA		

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001). Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000). RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000). Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001). cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
source

```

location/Qualifiers
1. .932
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="4833424007"
/sex="mixed"
/tissue_type="head"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 0 day neonate
head"
/notes="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'.
GAGAGAGGAAGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 100.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATCTCGATTAAATAATATCCCTCCCTCCCTCC
3']. cDNA was cloned into the XhoI and BamHI sites.
Vector: a modified pBluescript KS(+) after bulk excision
from Lambda FLX I."

ORIGIN

Query Match 52.0%; Score 623; DB 13; Length 932;
Best Local Similarity 83.9%; Pred. No. 1.9e-109;
Matches 726; Conservative 0; Mismatches 137; Indels 2; Gaps 2;

Qy 25 AAACAAGGAAAGTAGTACTGTGCTGAGAGACCATGGCAAAGAAATCTCCAGAGAAATG 84
Db 1 AAAGCTGGAAAGAGCACCGTGTGCTGGAGACCATGGCAAAGAAATCTCTCCAGAGAACTG 60

Qy 85 TGAAGACTGTACATCTTAAATCGAGAGCTTTTAAATCCAGAAATAATGTAATCACT 144
Db 61 TGAGGGCTGTACACATCTTAAATGAGAGAGCTGTGATATCTAAGATGATATGTTAAATCT 120

Qy 145 TAAGATTTTGCACTGGTGTGTTGGTATCCTGGCCCTAACTCTAATTTGCTGTTTGGGG 204
Db 121 GAAGATTTGTGACATGATGTTGGTATCCTGGCCTTAATCTAATTTGCTGTTTGGGG 180

Qy 205 GAGCAAGCACTTCTTGCCCGGAGGTACCCAAAAAGCCCTATGACATGGAGCACTTTTCTA 264
Db 181 GAGCAAAACACTTCTTGCCCGCATGTTTCCAGATAACCTATGATATGGAGCACTTTTCTA 240

Qy 265 CAGCAATGGAGAGAAGAGAGATTTCATCGAAATTCATCTGTGACAGCAACTCGAAT 324
Db 241 CAGCAACCGCAGAGAAGAGAGATTTCATGAAATTTGATCCCAATACCAACAGAAAT 300

```

ORIGIN

Query Match	52.0%;	Score 623;	DB 13;	Length 932;
Best Local Similarity	83.9%;	Pred. No. 1.9e-109;		
Matches 726;	Conservative	0;	Mismatches 137;	Indels 2;
Gaps	2;			
QY	25	AAACGAGGAAAGAGTACTGTGTCTCAGAGACCATGGCAAGAATCCTCCAGAGAAATG	84	
Db	1	AAAGCTGGGAAAGAGCACCGGTGTCTGGGAGACCATGGCAAGAATTCCTCCAGAGAACTG	60	
QY	85	TGAAGACTGTCACATTCCTAAATCAGAGAGCTTTTAAATCCAAGAAAATATGCTAAATCACT	144	
Db	61	TGAGGGCTGTGCACATTCCTAAATGCGAAGCTCTGATATCTAAGATGATATGTTAATTAAT	120	
QY	145	TAAGATTGTGGACTGGTGTGTGGTATCTTGGCCCTAACTTAATGTCTGTGTTTGGGG	204	
Db	121	GAAGATTGTGGACTAGTGTGTGGTATCCTGGCCTTAACCTAATGTGTCTGTGTTGGGG	180	
QY	205	GAGCAAGCACTTCTGGCGGAGGTACCCAAAAAGCCCTATCACATGGAGCACACTTTCTA	264	
Db	181	GAGCAAAACATTCTGGCCCGATGTATCCAGATAACCTATGATATGAGACACACTTTCTA	240	
QY	265	CAGCAATGGAGAGAAGAGATTTTACATGGAAATTTGATCCTGTGACAGCAACTGAAAT	324	
Db	241	CAGCAACGGCAGAGAAGAGATTTACATGGAAATTTGATCCCAATACCGAACAAGAAAT	300	


```
/db_xref="taxon:10090"
/clone="IMAGE:30295278"
/tissue type="embryonic limb, maxilla and mandible"
/lab host="DH10B (phage-resistant)"
/clone lib="NIH_MGC_136"
/notes=Vector; PCWV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Normalized, full-length enriched library from pool of
mouse embryonic limb, maxilla and mandible, embryonic day
17.5, 18.5 and newborn (mandible (5, 4 and 1 limb and jaw
equivalents from respective days). Cloned directionally,
oligo-dT primed (5'-GACTAGTCTAGATCGGCGGCGCC(T)15-3'.
Size selected for the >1kb fragments, average insert size
1.2 kb. Normalization to Cot 7.5. Tissue contributed by
David Rowe; library constructed by ResGen, Invitrogen
Corp. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 51.8%; Score 620.6; DB 14; Length 876;
Best Local Similarity 89.4%; Pred. No. 5.7e-109;
Matches 680; Conservative 0; Mismatches 79; Indels 2; Gaps 1;

QY 11 CTCTAGTCTCTCAAGCAAGGAAGAGTACTGTGTCTGAGAGACATGGCAAGAAT 70
Db 16 CTCTAGTCTCTCAAGCAAGGAAGAGACCGTGTGTCTGGGAGACATGGCAAGAAT 75
QY 71 CCTCCAGAGAAATTGTGAAGACTGTCAATCTTAATCAGAAGCTTTTAAATCCAAGAAA 130
Db 76 CCTCCAGAGAACTGTGAGGGCTGTCAATCTTAATCAGAAGCTCTGAATCTAAGAAG 135
QY 131 ATATGTAATCACTTAAGATTGTGGACTGGTGTGGTATCTCTGGCCCTAACTTAATT 190
Db 136 ATATGTAATCACTTAAGATTGTGGACTAGTGTGGTATCTCTGGCCCTTAATT 195
QY 191 GTCTCTTTTGGGGAGCAAGCACTTCTGCCCGAGGTACCCAAAAGCCTATGACATG 250
Db 196 GTCTCTTTTGGGGAGCAACACTTCTGCCCGAGGTATCCAGAAAACCTATGACATG 255
QY 251 GAGCACACTTCTACAGCAATGAGAGAGAAAGATTTCATGGAATTTGATCCCAT 310
Db 256 GAGCACACTTCTACAGCAACGGCAGAGAAAGATTTCATGGAATTTGATCCCAT 315
QY 311 ACCAGAACTGAATATTCAGACGGCAATGGCACTGATGAACATTTGGAAGTGCAGC 370
Db 316 ACCAGAACTGAATATTCAGAGTGAATTTGCAAGTGAATGGCACTGATGAACATTTGGAAGTGCAGC 375
QY 371 TTTAAAAACGATACACTGGCATCTACTTCTGTGGGTCTTCAAAAATGTTTATCAAAACT 430
Db 376 TTTAAAAATGGATACACTGGCATCTACTTGTAGGTCTTCAAAAATGCTTTATTAACCT 435
QY 431 CAGATTAAAGTATCTCTGAATTTCTGAACCAAGAGAGAAATAGATGAGATGAAGAA 490
Db 436 CAAATCAAAAGTATCTCTGAATTTCTGAACCAAGAGAGAAATAGATGAGATGAAGAA 495
QY 491 ATTACCACAACTTCTTCTGAACAGTCAAGTATTTGGTCCCGACAGAAAAGCCTATTGAA 550
Db 496 ATTACTCAACTTCTTCTGAACAGTCAAGTATTTGGTCCCGACAGAAAAGCCTATTGAA 555
QY 551 AACCGAGATTTCTTAAAAATTTCCAAAATTTCTGGAGATTTGTGATAACCTTGACCATGAT 610
Db 556 AACAGAGACTTCTGAAAAATTTCTAAAATTTCTGGAGATTTGCGAATGTCACATGATAC 615
QY 611 TGGATCAATCCCACTCTAATATCAGTTTCTGAGTTTACAGACTTTTGGAGGAGGAGAA 670
Db 616 TGGATCAATCCCACTCTAATAGCAGTTTTCAGAATTTACAGGACTTTTGGAGGAGGAGAA 675
QY 671 GATCTTCACTTTCTGCCCACGAAAAAAGGATTTGAACAAAATGAACAGTGGGTGG-- 728
Db 676 GATCTTCACTTTCTACCACTGAAAAAAGGATTTGACCAAGATGAGCAATGGGGGGT 735
QY 729 TCCCTCAAGTGAAGTAGAGAGACCCCGTCAGCCAGACAA 769
Db 736 CCGCCCAAGTGAAGTGGAGAAAGACCCCGCCACACAGAA 776
```

```
RESULT 16
CD512100 751 bp mRNA linear EST 06-JUN-2003
LOCUS AGENCOURT 14354032 NIH_MGC_186 Homo sapiens cDNA clone
DEFINITION IMAGE:30406681 5', mRNA sequence.
ACCESSION CD512100
VERSION CD512100.1 GI:31443818
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 751)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: gcapos-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
cDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM195 row: n column: 02
High quality sequence stop: 553.
FEATURES
Location/Qualifiers
1..751
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30406681"
/lab_host="NIH_MGC_186"
/clone_lib="NIH_MGC_186"
/notes="Organ: Pooled-Skin; Vector: pDNR-LIB; Site 1: SfII
(ggcatatggcc); Site 2: SfII (ggcgctctggcc); Library is
oligo-dT primed and directionally cloned. cDNA was
prepared from a pooled samples of tissues from Skin,
meninges, duramater, pia matter and choroid plexus.
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGCGGCACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library"
```

```
ORIGIN
Query Match 51.2%; Score 612.6; DB 14; Length 751;
Best Local Similarity 99.2%; Pred. No. 2e-107;
Matches 626; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 237 AAGCCATATGACATGGAGCACACTTTCTACAGCAATGGAGAGAGAGATTACATCG 296
Db 54 AGGCCATATGACATGGAGCACACTTTCTACAGCAATGGAGAGAGAGATTACATCG 113
QY 297 AAATTGATCTCTGTGACCAAGAACTGAATATTTCAGAGCGGAAATGGCACTGATGAACAT 356
Db 114 AAATTGATCTCTGTGACCAAGAACTGAATATTTCAGAGCGGAAATGGCACTGATGAACAT 173
QY 357 TGGAGTTCACGACTTTTAAAAACGATACACTGGCATCTACTTCGTGGGTCTTCAAAAT 416
Db 174 TGGAGTTCACGACTTTTAAAAACGATACACTGGCATCTACTTCGTGGGTCTTCAAAAT 233
QY 417 GTTTTATCAAAACTCAGATTAAAGTGATTCTTGTAATTTTCTGAACCAAGAGGAATAG 476
```

Db 234 GTTTTATCAAACTCAGATTAAAGTGATTCCTGAATTTCTGAACAGAGGAAATAG 293

QY 477 ATGAGATGAAGAAATACCAACTTCTTTGACAGTCAGTGAATTTGGTCCCGACAG 536

Db 294 ATGAGATGAAGAAATACCAACTTCTTTGACAGTCAGTGAATTTGGTCCCGACAG 353

QY 537 AAAAGCCTATTGAACCGGAGATTTCTTAAAAATTCMAAATTCGAGATTTGTGATA 596

Db 354 AAAAGCCTATTGAACCGGAGATTTCTTAAAAATTCMAAATTCGAGATTTGTGATA 413

QY 597 ACGTGACCATGATTGGATCAATCCACTCTAATATCACTGTTTCTGAGTTCAAGACTTTG 656

Db 414 ACGTGACCATGATTGGATCAATCCACTCTAATATCACTGTTTCTGAGTTCAAGACTTTG 473

QY 657 AGGAGGAGGAGAGATCTTCACTTCTGCAACGAAAAAAGGATTCGACAAAAATG 716

Db 474 AGGAGGAGGAGAGATCTTCACTTCTGCAACGAAAAAAGGATTCGACAAAAATG 533

QY 717 AACAGTGGTGGTCCCTCAAGTGAAGTAGAAGACCCGTCACGCCAGACAAAGCAAGTG 776

Db 534 AACAGTGGTGGTCCCTCAAGTGAAGTAGAAGACCCGTCACGCCAGACAAAGCAAGTG 593

QY 777 AGGAAGAACTTCCATAAATGACTATACGTAATGAAATGGAATGATTCATCCATGCTGG 836

Db 594 AGGAAGAACTTCCATAAATGACTATACGTAATGAAATGGAATGATTCATCCATGCTGG 653

QY 837 ATGAGAG-AGTTTATGTTGTTTACTGCC 866

Db 654 ATGAGAGAGATTTATGGTTGTTTACTGCC 684

RESULT 17

CB587917 867 bp mRNA linear EST 03-APR-2003

LOCUS AGENCOURT 12990010 NIH_MGC_136 Mus musculus cDNA clone

DEFINITION IMAGE:30292458 5', mRNA sequence.

CB587917

CB587917.1 GI:29505773

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 867)

NIH-MGC <http://mgs.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. David Rowe

cDNA Library Preparation: Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: NDAM336 row: f column: 19

High quality sequence stop: 605.

Location/Qualifiers

1. .867

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:30292458"

/tissue_type="embryonic limb, maxilla and mandible"

/lab_host="NIH_MGC_136"

/clone_lib="NIH_MGC_136"

/note="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI; Normalized, full-length enriched library from pool of mouse embryonic limb, maxilla and mandible, embryonic day 17.5, 18.5 and newborn (mandible (5, 4 and 1 limb and jaw equivalents from respective days). Cloned directionally, oligo-dT primed (5'-GACTAGTTCCTAGATCGCGCGCGCCCC(7)15-3'.

Size selected for the >1kb fragments, average insert size 1.2 kb. Normalization to Cot 7.5. Tissue contributed by David Rowe; library constructed by Resgen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ORIGIN

Query Match

Best Local Similarity 50.8%; Score 608.2; DB 14; Length 867;

Matches 646; Conservative 91.1%; Pred. No. 1.3e-106; Mismatches 63; Indels 0; Gaps 0;

QY 47 TGCTGAGAGACCATGGCAAGAAATCCCTCCAGAGAAATGTGAAGACTGTGCACATCTCTAAAT 106

Db 1 TGCTGGGAGACCATGGCAAAATCTCTCCAGAGAACTGTGAGGGCTGTGCACATCTCTAAAT 60

QY 107 GCAGAAAGCTTTTAAATCCAGAAAAATATGTAATATCACTTAAAGATTTGTGGAGCTGTGTTT 166

Db 61 GCAGAAAGCTCTGAAATCTTAAAGATATGTAATATCACTGAAAGATTTGTGGACTAGTGT 120

QY 167 GGTATCTCTGGCCCTAACTCTAAATTTGTCTGTTTGGGGGAGCAAGCACTTCTGGCCGGAG 226

Db 121 GGTATCTCTGGCCCTTAACTCTAAATTTGTCTGTTTGGGGGAGCAAAACATTTCTGGCCGGAG 180

QY 227 GTACCCAAAAAGCCTATGACATGGAGCACACTTTTCTACAGCAATGAGAGAAAGAAAG 286

Db 191 GTATCCAAAGAAACCTATGATGAGAGCACACTTTCTACAGCAACGCGGAGAGAAAGAAAG 240

QY 287 ATTTACATGGAATTTGATCTCTGTCACGAGCACTTTTCTACAGCAATGAGAGAAAGAAAG 346

Db 241 ATTTACATGGAATTTGATCTCTGTCACGAGCACTTTTCTACAGCAATGAGAGAAAGAAAG 300

QY 347 GATGAAACATTTGGAAGTGCAGCACTTTTAAACCGATACACTTGGCATCTACTTCTGGGT 406

Db 301 GATGAAACATTTGGAAGTGCAGCACTTTTAAACCGATACACTTGGCATCTACTTCTGGGT 360

QY 407 CTTCAAAAATGTTTATCAAACTCAGATTTAAAGTGAATCTCTGAAATTTCTCAACAGAA 466

Db 361 CTTCAAAAATGCTTTTATTAAGTCAATCAAGTGAATCTCTGAAATTTCTCAACAGAA 420

QY 467 GAGGAAATAGATGAGAAATGAGAAATTACCACAACTTTCTTTGAACAGTCACTGATTGG 526

Db 421 GAAGAAATAGATGAGAAATGAGAAATTTACTACAACTTTCTTTGAACAGTCACTGATTGG 480

QY 527 GTCCAGCAGAAAGCCCTATTGAAACCGAGATTTTCTTAAAAATTCCTCAAAATTCGGAG 586

Db 481 GTTCCCGCAGAAAGCCCTATTGAAACCGAGACTTCTCTGAAAAATTTCTAAAAATTCGGAG 540

QY 587 ATTTGTGATAAGCTGACCATGATTTGGATCAATCCCACTCTCTAATATCACTTTCTGAGTTA 646

Db 541 ATTTGGGATTAATGTGACCATGATTTGGATCAATCCCACTCTCTAATATCACTTTCTGAGAA 600

QY 647 CAAGACTTTTGAGAGAGGAGGAGAGATCTTCACTTTCTCTGCCAACGAAAAAAGGGATTT 706

Db 601 CAGGACTTTTGAGAGAGGAGGAGGATCTTCACTTTCTCTACCACTGAGAGAGAGGGGATT 660

QY 707 GAACAAAATGAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAGAGAGAGAGAGAGAGAGAG 755

Db 661 GACCAGAAATGAACAAATGGGTGGTCCCAAGTGAAGTAGAGAGAGAGAGAGAGAGAGAGAG 709

RESULT 18

BI454468

LOCUS 603171428F1 NCI_OGAP_Mam5 Mus musculus cDNA clone IMAGE:5250871 5', mRNA sequence.

DEFINITION BI454468

ACCESSION BI454468.1 GI:15245124

VERSION

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11633 row: h column: 08
High quality sequence start: 2
High quality sequence stop: 733.
High quality sequence stop: 733.

FEATURES
source
1. 744
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5250871"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP Mam5"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

ORIGIN
Query Match 48.4%; Score 579.6; DB 12; Length 744;
Best Local Similarity 90.7%; Pred. No. 4.2e-101;
Matches 640; Conservative 0; Mismatches 64; Indels 2; Gaps 2;

QY 11 CTCTCAGTCTCTCAAGACGAGGAAGAGTACTGTGTCTGAGAGACCATGGCAAGAAT 70
DB 35 CTCTCAGTCTCTCAAGACGAGGAAGAGACCGCTGTGTGGGAGACCATGGCAAGAAT 94
QY 71 CCTCCAGAGAAATGTGAAGACTGTCAATCTTAATGCAGAACTTTAATCCAGAAA 130
DB 95 CCTCCAGAGAACTGTGAGGGCTGTCAATCTTAATGCAGAACTTTAATCCAGAA 154
QY 131 ATATGTAATCACTTAAGATTTGTGACTGGTGTGGTATCTTGGCCCTAACTCTAATT 190
DB 155 ATATGTAATCACTGAAGATTTGTGACTAGTGTGGTATCTTGGCCCTAACTCTAATT 214
QY 191 GTCTCTGTTTGGGGAGCAGACATTTCTGCGGAGGTACCCAAAAGCCTATGACATG 250
DB 215 GTCTCTGTTTGGGGAGCAACACTTTCTGCGCGAGGTATCCAAGAAAACCTATGACATG 274
QY 251 GAGCACACTTTCTACAGCAATGAGAGAGAAAGAAATTTACATGGAATTTGATCTGTG 310
DB 275 GAGCACACTTTCTACAGCAACGCGAGAGAAAGAAATTTACATGGAATTTGATCCATA 334
QY 311 ACAGAACTGAAATATTCAGACGGAATGGCACTGATGAACATTTGGAAGTGCACGAC 370
DB 335 ACCAGAAACAGAAATATTCAGAAAGTGAATGGCACTGATGAACATTTGGAAGTGCAC 394
QY 371 TTTAAAACGGAATACACTGGCATCTACTTCTGGGGTCTTCAAAAATGTTTATCAAACT 430
DB 395 TTTAAAACGGAATACACTGGCATCTACTTCTGGGGTCTTCAAAAATGTTTATCAAACT 454
QY 431 CAGATTAAAGTGATTCCTGAAATTTCTGAACCAAGAGAGGAATATAGATGAGATGAAGA 490
DB 455 CAAATCAAGTGNATTCCTGAATTTCTGAACCAAGAGAGGAATATAGATGAGATGAAGA 514
QY 491 ATTAACCAACTTTCTTGAACAGTCAAGTGAATTTGGTCCCGAGCAAAAAGCCTATTGAA 550
DB 515 ATTAACCAACTTTCTTGAACAGTCAAGTGAATTTGGTCCCGAGCAAAAAGCCTATTGAA 574
QY 551 AACCGAGATTTCTTAAAATTCAAAATCTGGAGATTGTG-TGATAACGTGACCATGTA 609

DB 575 AACAGAGACTTCTCGAAAAATTTAAAAATTTGGAGATTTGCCGATATGTGACCATGTA 634
QY 610 TTGATCAATCCCACTCTPAA-TATCAGTTTCTGAGTTTACAGACTTTGAGGAGGAGGAG 668
DB 635 CTGGATCAATCCCACTCTTAATTAGCAGTTTCAGAATTTACAGACTTTTGGAGGAGCGGTG 694
QY 669 AAGATCTTCACTTCCCTCCCAACGAAAGAAAAGGATTGAACAAA 714
DB 695 AAGATCTTCACTTCTTCTACCAGTGAACAAAAGGGGATGATGCCAGA 740

RESULT 19
BF182340
LOCUS BF182340
DEFINITION 601804094F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4035214 5',
mRNA sequence.
ACCESSION BF182340
VERSION BF182340.1 GI:11060482
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 846)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9308 row: k column: 23
High quality sequence stop: 703.
High quality sequence stop: 703.

FEATURES
source
1. 846
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4035214"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP Mam5"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

ORIGIN
Query Match 48.4%; Score 579; DB 10; Length 846;
Best Local Similarity 88.2%; Pred. No. 5.3e-101;
Matches 675; Conservative 0; Mismatches 85; Indels 5; Gaps 4;

QY 11 CTCTCAGTCTCTCAAGCAAGGAAGAGTACTGTGTCTGAGAGACCATGGCAAGAAT 70
DB 32 CTCTCAGTCTCTCAAGCAAGGAAGAGACCGTGTGTGGGAGACCATGGCAAGAAT 91
QY 71 CCTCCAGAGAAATTTGAAGACTGTCACTTTAAATGCAGAGCTTTAAATCCAGAAA 130
DB 92 CCTCCAGAGAACTGTGAGGGCTGTCAATTCATAATGCAGAAAGCTCTGAAATCTAAGAAG 151
QY 131 ATATGTAATCACTTAAGATTTGTGACTGGTGTGGTATCTTGGCCCTAACTCTAATT 190
DB 152 ATATGTAATCACTGAAGATTTGTGACTAGTGTGGTATCTTGGCCCTAACTCTAATT 211

```
QY 191 GTCCTGTTTTGGGGAGCAACACTTCTGGCCGAGGTACCCAAAAAGCCTATGACATG 250
Db 212 GTCTCTGTTTTGGGGAGCAACACTTCTGGCCGAGGTATCCAGAAACCTATGACATG 271
QY 251 GAGCACACTTCTTACAGCAATGGAGAGAAAGAGATTTCATGGAATTTGATCTCTGTG 310
Db 272 GAGCACACTTCTTACAGCAACGGCGAGAAAGAGATTTCATGGAATTTGATCTCTGTG 331
QY 311 ACCGAACTGAAATATTCAGAGCGGAATGGCACTGATGAAATTCGGAAGTGCACGAC 370
Db 332 ACCGAACTGAAATATTCAGAGCGGAATGGCACTGATGAAATTCGGAAGTGCACGAC 391
QY 371 TTTAAAAACGGATACACTGGCATCTACTTCTGGGTCTTTCAAAAATGTTTTATCAAACT 430
Db 392 TTTAAAAATGATACACTGGCATCTACTTCTGGGTCTTTCAAAAATGTTTTATCAAACT 451
QY 431 CAGATTAAAGTGATTCCTGAAATTTCTGAAACAGAGAGGAATATGATGAGAAATGAAGA 490
Db 452 CAAATCAAAAGTATTCCTGAAATTTCTGAAACAGAGAGGAATATGATGAGAAATGAAGA 511
QY 491 ATTACCAACAATTTCTTTGAAACAGTCAGTGATTTGGTCCCGAGAGAAAGCCTATTGAA 550
Db 512 ATTACTACAA-TTTCTTTGAAACAGTCAGTGATTTGGTCCCGAGAGAAAGCCTATTGAA 570
QY 551 AACCAGAGATTTTCTTAAAAATTCCTGAAATTTCTGAGATTTGAT-AACGTGACCATGTA 609
Db 571 AACAGAGACTTCTGAAATTTCTTAAATTTCTGGAGATTTGGATTAATGTGACCATGTA 630
QY 610 TTGGATCAATCCACTCTTAATATCAGTTTCTGAGTTTCAAGACTTTGAGGAGAGGGAGA 669
Db 631 CTGGATCAATCCACTCTTAATATCAGTTTCTGAGTTTCAAGACTTTGAGGAGAGGGAGA 690
QY 670 AGATCTTCACTTCTCTGCAACGAAAAAAGGATTTGAACAAAAATGAACAGTGGGTGGT 729
Db 691 AGATCTTCACTTCTTACCAAGT-AAAAAGGGGATTTGACCAATGACCATGCGGGGGT 749
QY 730 CCTCAAGTGAAGTAGAAGAACCGTCAACGCGACGACAAAGCAAG 774
Db 750 CC--GCAGTGGGGGGGGGAAAAACCGCAACAGAAAGAGGAGG 792

RESULT 20
BG174435
LOCUS 60233425F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4457652 5',
DEFINITION mRNA sequence.
ACCESSION BG174435
VERSION BG174435.1 GI:12681138
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 958)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapps-f@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10254 row: m column: 13
High quality sequence start: 10
High quality sequence stop: 652.
Location/Qualifiers
1..958
/organism="Mus musculus"
/mol_type="mRNA"
```

```
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4457652"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam1"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

ORIGIN
Query Match 48.2%; Score 577; DB 12; Length 958;
Best Local Similarity 89.3%; Pred. No. 1.2e-100;
Matches 633; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

QY 41 ACTGTGCTCTCAGAGACCATGCAAGAAATCTCCAGAGAAATTTGTGAAGACTGTGCACATT 100
Db 2 ACCGTGTCTCGGAGACCATGCAAGAAATCTCCAGAGAAATTTGTGAAGACTGTGCACATT 61
QY 101 CTAAATGAGAGAGCTTTTAAATCCAGAAATAATGTAATCACTTAAGATTGTGGACTG 160
Db 62 CTAAATGAGAGAGCTCTGAAATCTAAGAAATATGTAATCACTTAAGATTGTGGACTG 121
QY 161 GTGTTGTTGATTCCTGCGCTTAATCTAATTTGCTGTTTTCGGGAGCAAGCACTTCTGG 220
Db 122 GTG-TTGGTATCTCGGCTTAACTCTAATTTGCTGTTTTCGGGAGCAAGCACTTCTGG 180
QY 221 CCGGAGGTACCCAAAAAGCCCTATGACATGGAGACACATTTCTCAGCAATGGAGAGAAG 280
Db 181 CCGGAGGTATCCAGAAACCTATGACATGGAGACACATTTCTCAGCAACGCGGAGAAG 240
QY 281 AAGAAGATTTTACATGGAATTTGATCTCTGTGACAGAACTGGAATATTGAGAGCGGAAT 340
Db 241 AAGAAGATTTTACATGGAATTTGATCTCTGTGACAGAACTGGAATATTGAGAGCGGAAT 300
QY 341 GGCACGTGATGAACATTTGGAAGTGCAGACTTTTAAACCGGATACACTGCATCTACTTC 400
Db 301 GGCACGTGATGAACATTTGGAAGTGCAGACTTTTAAACCGGATACACTGCATCTACTTC 360
QY 401 GTGGGTCTTCAAAAATGTTTATCAAACTCAGATTAAAGTGAATTCCTGAATTTCTGAA 460
Db 361 GTAGGTCTTCAAAAATGTTTATCAAACTCAGATTAAAGTGAATTCCTGAATTTCTGAA 420
QY 461 CCAGAGAGGAAATAGATGAGAAATGAAGAAATTAACAACTTTTCTTGAACAGTCAGTG 520
Db 421 CCAGAGGAAAGAAATAGATGAGAAATGAAGAAATTAACAACTTTTCTTGAACAGTCAGTG 480
QY 521 ATTGGGTCCAGCAGAAAGCCCTATTGAAACCGGAGATTTCTTAAATAATTCCTGAAAT 580
Db 481 ATTGGGTCCAGCAGAAAGCCCTATTGAAACCGGAGATTTCTTAAATAATTCCTGAAAT 540
QY 581 CTGAGATTTGTGATAACGTGACCATGTTGATGATCAATCCCACTCTTAATATCAGTTTCT 640
Db 541 CTGAGATTTGTGATAACGTGACCATGTTGATGATCAATCCCACTCTTAATATCAGTTTCT 600
QY 641 GAGTTTCAAGACTTTGAGGAGGAGGAGAGATCTTCACTTCTGCGCAACGAAAGAAAA 700
Db 601 GAATTACAGGACTTTGAGGAGGAGGAGAGATCTTCACTTCTGCGCAACGAGAACCCAG 660
QY 701 GGGATTGAACAAATGAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGA 749
Db 661 GGGATTGGCCAGAAATGAGCAATGGTGGTCCCGCAGCTGAGGTGAGAA 709

RESULT 21
CD103859 712 bp mRNA linear EST 15-MAY-2003
LOCUS AGENCOURT_14008022 NIH MGC_186 Homo sapiens cDNA clone
DEFINITION IMAGE:30372251 5', mRNA sequence.
ACCESSION CD103859
VERSION CD103859.1 GI:30757033
```

```

KEYWORDS      EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens

REFERENCE      1 (bases 1 to 712)
AUTHORS       NIH-MGC http://mgi.nci.nih.gov/.
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgabs-remail.nih.gov
              Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
              cDNA Library Preparation: CLONTECH Laboratories, Inc.
              DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: NDCM164 row: c column: 12
              High quality sequence stop: 494.
              Location/Qualifiers
                1..712
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:30372251"
                  /lab_host="DH10B (T1 phage-resistant)"
                  /note="Organ: Pooled-Skin; Vector: pDNR-LIB; Site 1: Sfil
                    (ggccattatggcc); Site 2: Sfil (ggcgctcctggcc); Library is
                    oligo-dT primed and directionally cloned. cDNA was
                    prepared from a pooled samples of tissues from skin,
                    meninges, duramater, pia matter and choroid plexus.
                    and 3' adaptors were used in cloning as follows: 5'
                    adaptor sequence: 5'-CAGCGCATTTATGCGC-3' and 3' adaptor
                    sequence: 5'-ATTCTAGAGCGCGGCGGCGCATG-dT(30)BN-3'
                    (where B = A, C, or G and N = A, C, G, or T). Average
                    insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies
                    contained inserts by PCR. This library was enriched for
                    full-length clones and was constructed by Clontech
                    Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                    Library"

FEATURES       source
              Location/Qualifiers
                1..679
                  /organism="Mus musculus"
                  /mol_type="mRNA"
                  /strain="C57BL/6J"
                  /db_xref="taxon:10090"
                  /clone="IMAGE:5148842"
                  /tissue_type="tumor, gross tissue"
                  /dev_stage="7 months"
                  /lab_host="DH10B"
                  /clone_lib="NCI CGAP Mam5"
                  /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sall;
                    Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                    Library constructed by Life Technologies. Investigators
                    providing samples: Lothar Hennighausen/Robin Humphreys,
                    NIH"

ORIGIN
Query Match      48.0%; Score 574; DB 14; Length 712;
Best Local Similarity 98.8%; Pred. No. 5.1e-100;
Matches 589; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 237 AAGCCTATGACATGGAGCACACTTCTACAGCAATGGAGAGAGAGAGATTTACATGG 296
DB 58 AGCCTATGACATGGAGCACACTTCTACAGCAATGGAGAGAGAGAGATTTACATGG 117
QY 297 AAATTGATCTGTGACAGAACTGAATAATTTCAGAGCGGAATGGCACTGATGAACAT 356
DB 118 AAATTGATCTGTGACAGAACTGAATAATTTCAGAGCGGAATGGCACTGATGAACAT 177
QY 357 TCGAAGTGCAGCACTTTAAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAAT 416
DB 178 TCGAAGTGCAGCACTTTAAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAAT 237
QY 417 GTTTATCAAACTCAGATTAAAGTGAATTCCTGAATTTTCGAACAGAGAGAAATAG 476
DB 238 GTTTATCAAACTCAGATTAAAGTGAATTCCTGAATTTTCGAACAGAGAGAAATAG 297
QY 477 ATGAGAAATGAAGAAATTAACACAACTTTCTTTCAGACAGTCAGTGAATTCGGTCCGACG 536
DB 298 ATGAGAAATGAAGAAATTAACACAACTTTCTTTCAGACAGTCAGTGAATTCGGTCCGACG 357
QY 537 AAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCAAAATTTCTGGAGATTTTGATA 596
DB 358 AAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCAAAATTTCTGGAGATTTTGATA 417
QY 597 ACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTTACAAGACTTTG 656

```

```

Db 418 ACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTTACAAGACTTTG 477
QY 657 AGGAGGAGGAGAGAGATCTTCACTTTCTCCCAACGAAAAAAGGATTTGAACAAATG 716
DB 478 AGGAGGAGGAGAGAGATCTTCACTTTCTCCCAACGAAAAAAGGATTTGAACAAATG 537
QY 717 ACAGTGGGTGGTCCCTCAAGTGAAGTAGAAGACCCGTCAAGCCAGACAGCAAGTGG 776
DB 538 AACAGTGGGTGGTCCCTCAAGTGAAGTAGAAGACCCGTCAAGCCAGACAGCAAGTGG 597
QY 777 AGGAGAACTTCCAATAAATGACTACTGAAAA--TGAATAGATTTGATCCCA 830
DB 598 ATGAAGAACTTCCAATAAATGACTACTGAAAAATGGAATAGATTTGATCCCA 653

RESULT 22
BI248748
LOCUS      679 bp      mRNA      linear      EST 17-JUL-2001
DEFINITION Mus musculus cDNA clone IMAGE:5148842 5',
            mRNA sequence.
ACCESSION  BI248748
VERSION    BI248748.1 GI:14795446
KEYWORDS  EST.
SOURCE     Mus musculus (house mouse)
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11367 row: m column: 03
High quality sequence stop: 677.
Location/Qualifiers
  1..679
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="C57BL/6J"
    /db_xref="taxon:10090"
    /clone="IMAGE:5148842"
    /tissue_type="tumor, gross tissue"
    /dev_stage="7 months"
    /lab_host="DH10B"
    /clone_lib="NCI CGAP Mam5"
    /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sall;
      Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
      Library constructed by Life Technologies. Investigators
      providing samples: Lothar Hennighausen/Robin Humphreys,
      NIH"

ORIGIN
Query Match      46.2%; Score 553.2; DB 12; Length 679;
Best Local Similarity 91.7%; Pred. No. 5e-96;
Matches 607; Conservative 0; Mismatches 53; Indels 2; Gaps 2;

QY 11 CTCTCAGTCTCTCAAGCAAGGAGAGAGTACTGTCTGCTGAGAGACCATGCAAGAAT 70
DB 19 CTCTCAGTCTCTCAAGC-AGGAAGAGACACCGTGTCTGGGAGACCATGCAAGAAT 77
QY 71 CTCCAGAGAAATTTGAGAGACTGTGCACATTCTTAAATGCAAGAGCTTTTAAATCCAGAA 130
DB 78 CTCCAGAGAACTGTGAGGGCTGTGCACATTCTTAAATGCAAGAGCTCTGAAATCTAAGAAG 137

```

```

Qy 131 ATATGTAATCATTACATTTGGAAGTGGTGTGTTGGTATCCTGGCCCTAACTTAAT 190
    |||||
Db 138 ATATGTAATCATTACGAAGATTGTGGACTAGTGTGGTATCCTGGCCCTTAAT 197
    |||||
Qy 191 GTCCTGTTTGGGGAGCAGACACTTCTGGCCGGAGGTACCACAAAAGCCTATGACATG 250
    |||||
Db 198 GTCCTGTTTGGGGAGCAGACACTTCTGGCCGGAGGTATCCAGAAAACCTATGACATG 256
    |||||
Qy 251 GAGCACACTTCTCAGCAATGGAGAGAGAGATTAATGATGAATGATCCTCTG 310
    |||||
Db 257 GAGCACACTTCTCAGCAACGGCAGAGAGAGATTTATGATGAATGATCCCAT 316
    |||||
Qy 311 ACCAGAACTGAATATTCAGAGCGGAATGSCACTGATGAACATGGAAGTCAGCAG 370
    |||||
Db 317 ACCAGAACTGAATATTCAGAGCGGAATGGAATGCACTGATGAACATGGAAGTCAGCAG 376
    |||||
Qy 371 TTTAAAAACGGATACACTGGCACTACTTCTGGGTCTTCAAAAATGTTTATCAAACT 430
    |||||
Db 377 TTTAAAAACGGATACACTGGCACTACTTCTGGGTCTTCAAAAATGTTTATCAAACT 436
    |||||
Qy 431 CAGATTAAGTGAATCCTGCAATTTCTGACCAAGAGAGGAAATAGATGAGAAATGAAGA 490
    |||||
Db 437 CAAATCAAAAGTGAATCCTGCAATTTCTGACCAAGAGAGGAAATAGATGAGAAATGAAGA 496
    |||||
Qy 491 ATTACCAAACTTCTTTTGAACAGTCAGTGAATTTGGTCCCGCAGAGAAAGCCTATTGAA 550
    |||||
Db 497 ATTACTACAACTTCTTTTGAACAGTCAGTGAATTTGGTCCCGCAGAGAAAGCCTATTGAA 556
    |||||
Qy 551 AACCAGAGATTTCTTAAAAATTCAAAATCTTGAGATTTGTGATTAACGTGACCATGTAT 610
    |||||
Db 557 AACAGAGACTTCTCGAAAAATCTCAAAATCTTGAGATTTGCGATATGTGACCATGTAT 616
    |||||
Qy 611 TGGATCAATCCCACTCTAATATACAGTTCTGACCTTACAGACTTTCAGAGAGAGGAGAA 670
    |||||
Db 617 TGGATCAATCCCACTCTAATATAGCAGTTTCAGAAATTAACAGACTTTCAGAGAGAGGAGAA 676
    |||||
Qy 671 GA 672
    ||
Db 677 GA 678

RESULT 23
AV593197 600 bp mRNA linear EST 27-NOV-2001
LOCUS AV593197 Bos taurus cartilage fetus Bos taurus cDNA clone
DEFINITION E1CA006A12 5', mRNA sequence.
ACCESSION AV593197
VERSION AV593197.1 GI:9708354
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
Takesuga, A., Hirotsune, S., Itoh, R., Jitohzono, A., Suzuki, H., Aso, H.
1 (bases 1 to 600)
Takasuga, A., Hirotsune, S., Itoh, R., Jitohzono, A., Suzuki, H., Aso, H.
Establishment of a high throughput EST sequencing system using
poly(A) tail-removed cDNA libraries and determination of 36,000
bovine ESTs
Nucleic Acids Res. 29 (22), E108 (2001)
21570554
11713328
Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@cccoo.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
Location/Qualifiers

FEATURES

```

```

1. 600
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="E1CA006A12"
/tissue_type="cartilage"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_lib="Bos taurus cartilage fetus"
/notes="Vector: pZ11; Site_1: SalI; Site_2: NotI; Poly A
was deleted from a NotI site"

ORIGIN
Query Match 44.4%; Score 531.8; DB 9; Length 600;
Best Local Similarity 93.7%; Pred. No. 6.5e-92; Indels 0; Gaps 0;
Matches 554; Conservative

Qy 24 CAAACCAAGGAAGAGTACTGTGCTCGAGAGACCATGGCAAGAAATCTCCAGAGAAAT 83
    |||||
Db 10 CGAAGCAGGAAGAGAGTGTGCTGCGCGAGAGATCATGGCAAGAAATCTCCAGAGAACT 69
    |||||
Qy 84 GTGAAGACTGTCAATTTCTAAATGAGAGAGCTTTTAAATCCAGAAATATGTAATCAC 143
    |||||
Db 70 GTGAGACTGTCAATTTCTAAATGAGAGAGCTTTTAAATCCAGAAATATGTAATCAC 129
    |||||
Qy 144 TTAAGATTTGTGACTGGTGTGTTGGTATCCTGCGCCCTAACTCTAAATGTCCTGTTTGGG 203
    |||||
Db 130 TTAAGATTTGTGAGTTGGTATTTGGTATCCTGCGCTTAACTCTAAATGTCCTGTTTGGG 189
    |||||
Qy 204 GGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCCTATGACATGGAGCAGACATTTCT 263
    |||||
Db 190 GGAGTAAGCACTTCTGGCCCTGAGACACCCCAAAAAACATATGACATGGAGCAGACATTTCT 249
    |||||
Qy 264 ACAGCAATGGAGAGAGAGAGATTTATGATGAATTTGATCTCTGTGACGAGAACTGAA 323
    |||||
Db 250 ACAGCAATGGAGAGAGAGAGATTTATGATGAATTTGATCTCTGTGACGAGAACTGAA 309
    |||||
Qy 324 TATTCAGAGCGGAATGGCACTGATGAACATTTGGAAGTGACAGACTTTAAAAAACGGAT 383
    |||||
Db 310 TATTCAGAGCGGAATGGCACTGATGAACATTTGGAAGTGACAGACTTTAAAAAACGGAT 369
    |||||
Qy 384 ACATCGCATCTACTTCTGGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGA 443
    |||||
Db 370 ACATCGCATCTACTTCTGGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGA 429
    |||||
Qy 444 TTCTGAAATTTCTGAACCAAGAGAGAAATAGATGAGAAATGAGAAATTAACCAAACTT 503
    |||||
Db 430 TTCTGAAATTTCTGAACCAAGAGAGAAATAGATGAGAAATGAGAAATTAACCAAACTT 489
    |||||
Qy 504 TCTTTGACAGCTCAGTGAATTTGGTCCCGCAGAGAAAAGCCCTATTGAAAACCGAGATTTC 563
    |||||
Db 490 TCTTTGACAGCTCAGTGAATTTGGTCCCGCAGAGAAAAGCCCTATTGAAAACCGAGATTTC 549
    |||||
Qy 564 TTAATAATTCAAAATTTCTGAGATTTGTGATTAACCGTGAACATGATTGGA 614
    |||||
Db 550 TTAATAATTCAAAATTTCTGAGATTTGTGATTAACCGTGAACATGATTGGA 600
    |||||

RESULT 24
CD773806 639 bp mRNA linear EST 02-JUL-2003
LOCUS CD773806
DEFINITION AGENCOURT 14719761 NIH MGC 190 Mus musculus cDNA clone
IMAGE:30504556 5', mRNA sequence.
ACCESSION CD773806
VERSION CD773806.1 GI:32432308
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 639)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

```


JOURNAL
COMMENT

Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Yoshihiko Yamada, Takashi Nakamura, NIDCR
CDNA Library Preparation: Clontech Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM232 row: d column: 05
High quality sequence stop: 529.
Location/Qualifiers
1. .639
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 190"
/note="Organ: Pooled - Molar: Vector: pDNR-LIB; Site 1:
SfiI (ggccattatggcc); Site 2: SfiI (ggccctcctggcc);
Non-normalized full-length enriched library 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCCATTCATGGC-3' and 3' adaptor
sequence: 5'-ATTCAGAGCGCGCGGCACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.71
kb (range 0.5-3.0 kb). 15/15 colonies contained inserts by
PCR. This library was enriched for full-length clones and
was constructed by Clontech Laboratories (Palo Alto, CA)
Corp."

FEATURES
source

RESULT 25
BI694698
LOCUS
DEFINITION
603347638F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5375361 5',
mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BI694698 735 bp mRNA linear EST 18-SEP-2001
BI694698
BI694698.1 GI:15657327
EST.
Mus musculus (house mouse)
Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 735)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1955 row: k column: 10
High quality sequence start: 8
High quality sequence stop: 732.
Location/Qualifiers

FEATURES
source

1. .735
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5375361"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam2"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sall;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 41.8%; Score 500.6; DB 12; Length 735;
Best Local Similarity 88.7%; Pred. No. 5.7e-86;
Matches 611; Conservative 0; Mismatches 69; Indels 9; Gaps 6;
QY 246 ACATGGAGCACATTCTCTACA--GCANTGGAGAGAGAGAGATTGA--CATGGAAATTG 302
DB 47 ACATGGAGCACATTCTCTACATGCAACGGCGGAGAGAGAGATTACCATGGAATTG 106
QY 303 ATCTCTGTACCCAGAACTGAAATATTACAGAGCGGAATCGCACTGATGAAACATTGGAG 362
DB 107 ATCCCATACCGAAGACAGAAATATTTCAGAAATGGCAATGGCACTGATGAAACATTGGAAG 166
QY 363 TCACGACACTTTAAAAACGGATACACTGGCATCTACTTGTGGTCTTCAAAAATGTTTTA 422
DB 167 TCCATGACTTTAAAAATGAATACATGGCATCTACTTTGTAGGTCTTCAAAAATGCTTTA 226

ORIGIN

Query Match 44.3%; Score 530.4; DB 14; Length 639;
Best Local Similarity 91.5%; Pred. No. 1.2e-91;
Matches 572; Conservative 0; Mismatches 52; Indels 1; Gaps 1;
QY 20 CTCTCAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAAGAAATCCTCCAGAG 79
DB 7 CTCTCAAGCAAGGAAAGAGACCGTGTCTGTGGAGACCATGGCAAGAAATCCTCCAGAG 66
QY 80 AATTGTGAAGACTGTACATTCTTAATTCAGAACTTTTAATCCAAAGAAATATGTAAA 139
DB 67 AACTGTGAGGGCTGTCACTTCTTAATTCAGAACTCTGAAATCTAAGAGATATGTAAA 126
QY 140 TCACTTAAGATTTTGTGACTGGTGTGGTATCTCTGGCCCTTAACCTTAATGTCTCTTT 199
DB 127 TCACTGAAGATTGTGGACTAGTGTGGTATCTCTGGCCCTTAACCTTAATGTCTCTTT 186
QY 200 TGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCCTATGATGAGGACACT 259
DB 187 TGGGGGAGCAACACTTCTGGCCGGAGGTATCCAGAAACCTATGATGAGGACACT 246
QY 260 TTCTACAGCAATGGAGAGAGAAAGATTACATGGAAATTTGATCTGTGACCAAGACT 319
DB 247 TTCTACAGCAACCGCGAGAGAGAAAGATTACATGGAAATTTGATCCATAACAGACACA 306
QY 320 GAAATATTACAGAGCGGAATGGCACTGATGAACATTTGGAAGTGCACGCTTTAAAAAC 379
DB 307 GAAATATTGAAAGTGAATGGCACTGATGAACATTTGGAAGTGCACGCTTTAAAAAT 366
QY 380 GGATACACTGGCACTACTTCTGTGGTCTTCAAAAATGTTTATCAAACTCAGATTAAA 439
DB 367 GGATACACTGGCACTACTTGTAGTCTTCAAAAATGCTTTTAAAACTCAATCAAA 426
QY 440 GTGATTCCTGAATTTCTGAACAGAGAGGAATATAGATGAGAAATGAAGAAATACCAACA 499
DB 427 GTGATTCCTGAATTTCTGAACAGAGGAAGAAATATAGATGAGAAATGAAGAAATACCAACA 486
QY 500 ACTTCTTTTGAACAGTCAGTGATTGGGTCCCGAGCAGAAAAAGCCTATTGAAAAACCGAGAT 559

QY 423 TCRAAACTCAGATTAAAGTGATCTCTGAATTTCTGACCAAGAGAGAAATAGATGAGA 482
Db 227 TTAARACTCAAAATCAAGTGATCTCTGAATTTCTGACCAAGAGAGAAATAGATGAGA 286
QY 483 ATGAAGAAATATACCAACACTTTCTTTGAACAGTCTAGTGAATTTGGTCCGACAGAAAAAGC 542
Db 287 ATGAAGAAATATACCAACACTTTCTTTGAACAGTCTAGTGAATTTGGTCCGACAGAAAAAGC 346
QY 543 CTATTGAAAAACGAGATTTCTTTAAAAATTCAAAATTTCTGAGATTTGTGATACGTGA 602
Db 347 CTATTGAAAAACGAGATTTCTTTAAAAATTTCTGAGATTTGTGATACGTGA 406
QY 603 CCATGATTGGATCAATCCCACTCTAAATACATTTCTTCTGAGTTACAAGACTTTTGAGGAGG 662
Db 407 CCATGATTGGATCAATCCCACTCTAAATAGCAGTTTTCAGAAATACAGACTTTTGAGGAGG 466
QY 663 AGGGAAGAAGATCTTCACTTTCTTCCCAACGAAAAAAGGGATTTGAACAA--AATGAACA 720
Db 467 ACGGTGAAGATCTTCACTTTCTTCCCAACGAAAAAAGGGATTTGAACCAATGAGCA 526
QY 721 GTGGGTGGTCCCTCAAGTGAAGTAGA--GAAGACCCGTCACGCCAGACAAGCAAGTGAG 778
Db 527 ATGGGTGGTCCGCAAGTGAAGTGGACCGAAGACCCGCCACACAGCAAGCAAGCGAG 586
QY 779 GAAGACCTTCCATAATCACTATACATAATTTGAAAT--GGAATAGAAATTTGATCCCATGCTGGA 837
Db 587 GAAGACCTTCCATAATCACTATACATAATTTGAAATTTGAAATTTGATCCCATGCTGGA 646
QY 838 TGAGAGAGGTATTGTTGTTATTTACTGCGCTGAGGCAACCGCTATTTGCGCGCGCTCTG 897
Db 647 TGAGAGAGGTATTGTTGTTATTTACTGCTGCTGAGGCAACCGTACTTGGCGG--CGTGTCTG 705
QY 898 TGAACCTTTACTAGGCTACTACCCATATC 926
Db 706 TGAACCTTTACTAGGCTACTACCCATACC 734

RESULT 26
BI535437 557 bp mRNA linear EST 30-AUG-2001
LOCUS 398969 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BI535437
ACCESSION BI535437
VERSION BI535437.1 GI:15376545
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 557)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
Bennett,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A.,
Chitko-McKown,C.G., Perte,G., Holt,I., Karamycheva,S., Liang,F.,
Quackenbush,J. and Keele,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
11282978
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980504.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCAGCAGC
Plate: 127 row: C column: 3

Seq primer: ATTTAGGTGACACTATAG.
Location/Qualifiers
1..557
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4BOV"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 20 and day 40
embryos."

FEATURES
source

ORIGIN
Query Match 41.7%; Score 498.6; DB 12; Length 557;
Best Local Similarity 93.9%; Pred. No. 1.5e-85;
Matches 519; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 1 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGCAAGAGTACTGTGTCTGAGAGACCAT 60
Db 5 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGCAAGAGTACTGTGTCTGAGAGATCAT 64
QY 61 GGCAAGAAATCCCTCCAGAGAAATTTGAGAGACTGTCACTTCTAAATGCAAGACTTTTAA 120
Db 65 GGCAAGAAATCCCTCCAGAGAACTGTGAGGACTGTCACTTCTAAATGCAAGACTTTTAA 124
QY 121 ATCCAGAAAAATATGTAATCACTTAAAGATTTGTGAGACTGTGTTGGTATCTCGGCCCT 180
Db 125 ATCCAGAAAGATATGTAATTTACTTAAAGATTTGTGAGATTTGGTATCTCGGCCCT 184
QY 181 AACTCTAATTTGCTCTGTTTGGGGGAGCAAGCACTTCTGCGCGAGGTACCCAAAAAAGC 240
Db 185 AACTCTAATTTGCTCTGTTTGGGGGAGTAAAGCACTTCTGCGCTGAGACACCCAAAAAAGC 244
QY 241 CTATGACATGGAGCACACATTTCTACAGCAATGAGAGAGAGAAAGATTTACATGGAAT 300
Db 245 ATATGACATGGAGCACACATTTCTACAGCAATGAGAGAGAGAAAGATTTACATGGAAT 304
QY 301 TGATCTGTGACAGAACTGAAATATTCAGAACGGGAAATGSCACTGATCAAAACATTGGA 360
Db 305 TGATCCCATTTACCAAACTGAAATATTCAGAAAGTGGAAATGGCACTGATGAACATTGGA 364
QY 361 AGTGACAGCTTTAAACACGATACACTGTCATCTCTGCTGGGCTTCAAAAAATGTTT 420
Db 365 AGTACATGACTTTAAAAATGGATACATGCACTTACTTTGTAGTCTTCAAAAAATGTT 424
QY 421 TATCAAACTCAGATTAAGTGAATTTCTGAATTTTCTGAACCAAGAGAGAAATAGATGA 480
Db 425 CATCAAACTCAGATTAAGTGAATTTCTGAATTTTCTGAACCAAGAGAGAAATAGATGA 484
QY 481 GAATGAAGAAATTACCAAACTTTCTTTGAACAGTCACTGATTTGGGTCCGACAGAAAA 540
Db 485 GAATGAAGAAATTACCAAACTTTCTTTGAACAGTCACTGATTTGGGTCCGACAGAAAA 544
QY 541 GCCTATTGAAAC 553
Db 545 ACCTATTGAAAC 557

RESULT 27
AW743952 588 bp mRNA linear EST 27-APR-2000
LOCUS ur24h02.y1 Soares mouse NMHP Mus musculus cDNA clone IMAGE:3025299
DEFINITION 5', similar to TR:Q9YI63 Q9YI63 CHONDROMODULIN-1 PRECURSOR. ; mRNA
sequence.
ACCESSION AW743952
VERSION AW743952.1 GI:7655775
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 588)

AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information. MGI:1054007 Seq primer: -40RP from Gibco High quality sequence set: 486.

FEATURES
source

```

source
1. .588
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3025299"
/lab_host="PH10B (phage-resistant)"
/clone_lib="Soares mouse NMBP"
/note="Organ: bone (pooled); Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site.1: NotI;
Site.2: EcoRI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
TGTTCACCATCTGACGTAGCGGCGCCGACCGCTTTTTTTTTTTT 3'];
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT7T3 vector. Library
constructed and normalized by Bento Soares and M.Fátima
Bonaldo."

```

ORIGIN

Query Match	41.3%;	Score 494;	DB 10;	Length 586;
Best Local Similarity	91.3%;	Pred. No. 1.1e-84;		
Matches 524;	Conservative 0;	Mismatches 50;	Indels 0;	Gaps 0;
167	GGTATCTCTGGCCCTAACTCTAATGTGCTGTGTTTGGGGGAGCAAGCACTTCTGCGCGGAG	226		
15	GGTATCTCTGGCCCTAACTCTAATGTGCTGTGTTTGGGGGAGCAAACTTCTGCGCGGAG	74		
227	GTATCCCAAAAAGCCCTATGACATGGAGCACTTTTCTACAGCAATGAGAGAAAGAAG	286		
75	GTATCCCAAAAAGCCCTATGACATGGAGCACTTTTCTACAGCAACGCGGAGAGAGAAG	134		
287	ATTTACATGGAAATTTGATCCTCTGTCGACCACTGTAATTTTCAAGACGGGAATGGCACT	346		
135	ATTTACATGGAAATTTGATCCCAATACCCAGACAGAATATTTAGAAGTGGAAATGGCACT	194		
347	GATGAAACATTTGGAAGTGCA CGACTTTTAAAAACGGATACACTGGCACTCTACTTCGTGGGT	406		
195	GATGAAACATTTGGAAGTGCACTGACTTTAAAAATGGATACACTGGCACTCTACTTCGTAGGT	254		
407	CTTCAAAAATGTTTTATCAAAACTCAGATTTAAAGTGATTCCTGTAATTTTCTGAACCAAG	466		
255	CTTCAAAAATGCTTTATTAAAACTCAATCAAAAGTATTCCTGTAATTTTCTGAACCAAG	314		
467	GAGAAATAGATGAGATGAAGAAATTTACCAACACTTTTCTTTGAACAGTCAGTGAATTCGG	526		
315	GAGAAATAGATGAGATGAAGAAATTTACTACAACTTTCTTTGAACAGTCAGTGAATTCGG	374		
527	GTCCCGACAGAAAAAGCCTATTGAAAAACCGAGATTTCTTAAAAATTTCCAAATTTCTGGAG	586		
375	GTTCCTCCGACAGAAAAAGCCTATTGAAAAACAGAGACTTCCTGAAAAAATTTCAAAAATCTCGGAG	434		
587	ATTTGTGATACGTGACCAATGATTTGGATCAATCCCACTCTAATATCATAGTTTCTCAGTTTA	646		
435	ATTTTGGCATATATGTGACCATCTACTGGATCAATCCCACTCTAATATAGAGTTTCAGAAATTA	494		
647	CAAGACTTTGAGGAGGAGGAGAGACTTTCACATTTCTGCGCAACGAAAAAAGAGGATTT	706		
495	CAGGACTTTGAGGAGGAGCGGTGAAGATCTTCACATTTCTCTACAGTGAAAAAAAGGGGATTT	554		
707	GAACAAATGAAACAGTGGGTGGTCCCTCAAGTGA	740		

Db 555 GACCAGAAATGAGCAATGGGTGGTCCCGCAAGTGA 588

RESULT 28

BI249999	BI249999	6029524.F1	NCI_CGAP_Mam5	Mus musculus	cdna clone	linear	EST 17-JUL-2003
LOCUS							
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							

ORGANISM

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 798)	NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LUAM1373 row: a column: 12 High quality sequence stop: 796.		

FEATURES

```

1. .798
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5150867"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam5"
/note="Organ: mammary; Vector: pCMV
Site 2: NotI; Cloned unidirectional
Library constructed by Life Techno
providing samples: Lothar Hennighaus
NIH"
```

ORIGIN

Query Match	41.2%;	Score 493.6;	DB 12;	Length 798;
Best Local Similarity	91.5%;	Pred. No. 1.2e-84;		
Matches 567;	Conservative 0;	Mismatches 49;	Indels 4;	Gaps 4;
QY	11	CTCTCAGTCTCTCAAGCAAGCAAGAGTACTGTGTCTGTGAGAGCCATGGCCAAAGAAT	70	
Db	180	CTCTCAGTCTCTCTCAAGCAGGGAAGAGACCGTGTCTGTGGAGACCATGGCCAAAGAAT	239	
QY	71	CTTCCAGAGAAATTGTGAAGACTGTGCATTTCTTAATTCGAGAGCTTTTAAATCCCAAGAA	130	
Db	240	CTTCCAGAGAACTGTGAGGGCTGTGCATTTCTTAATTCGAGAGCTCTGAATCTAAGAAG	299	
QY	131	ATATGTAAATCATCTTAAGATTTTGTGGACTGGTGTTTGGTATCTGTGCCCTTAACCTCAATT	190	
Db	300	ATATGTAAATCATCTGAAGATTTTGTGGACTAGTG-TTGGTATCTGTGCCCTTAACCTCAATT	358	
QY	191	GTCCCTGTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCCTATGACATG	250	
Db	359	GTCCCTGTTTTGGGGGAGCAAAACACTTCTGGCCCGAGGTATCCAAGAAAAACCTATGACATG	418	
QY	251	GAGCACACTTTTCTACAGCAATTCGAGAGAAGAAGATTTTACATGAAATTCATCTGTG	310	
Db	419	GAGCACACTTTCTACAGCAACGGCGAGAGAGAAGATTTTACATGAAATTCATCCCAATA	478	
QY	311	ACCAGAACTGAATATTCAGAAAGCGGAATGGCACTGATGAACCAATTGGAACTGCACGAC	370	

```

Db      479 ACCGAGACAGAAAATTTCAGAAAGTGGAAATGGCACTGATGAAACATTGGAAAGTCCATGAC 538
QY      371 TTTAAAAACGGATACACTGCGATCTACTTCCTGGGCTCTTCAAAAATGTTTATCAAAACT 430
Db      539 TTTAAATGATACACTGCGATCTACTTCCTGGGCTCTTCAAAAATGTTTATCAAAACT 598
QY      431 CAGATTAAGTGAATTCCTGAATTTCTGAACCCAGAGAGGAAATAGATGAGAAATGAAGAA 490
Db      599 CAAATCAAAAGTGAATTCCTGAATTTCTGAACCCAGAGAGGAAATAGATGAGAAATGAAGAA 658
QY      491 ATTACCAACACTTCTTTTG-AACAGTCAGTGAATTTGGGTCCTCCAGCAG-AAAAGCCCTATTG 548
Db      659 ATTACTCAACACTTCTTTTGAAACAGTCAGTGAATTTGGGTCCTCCAGAGAAAGCCCTATTG 718
QY      549 AAAACCGAGATTTTCT-TAAAAATTCAAAATTCGGAGATTTTGATGAACCTGACCATG 607
Db      719 AAAACAGAGACTTCTCGAAAAAATTTCTAAAATTTCTGGAGATTTGGGATAATGTGACCATG 778
QY      608 TATTGATCAATCCCACTCT 627
Db      779 TACTGGATCAATCCCACTCT 798

RESULT 29
CB440637 610 bp mRNA linear EST 25-MAR-2003
LOCUS 690932 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION CB440637
VERSION CB440637.1 GI:29226859
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 610)
Smith, J.F.L., Roberts, A.J., Ehternkamp, S.E., Chitko-McKown, C.G.,
Wray, J.E. and Keefe, J.W.
A second set of bovine ESTs from pooled-tissue normalized libraries
Unpublished (2003)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: F018047 row: F column: 8
Seq primer: GTAATACGACTCACTATAGG.
FEATURES
Location/Qualifiers
1..610
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."
ORIGIN
Query Match 40.6%; Score 485.8; DB 14; Length 610;
Best Local Similarity 90.2%; Pred. No. 4.1e-83;
Matches 544; Conservative 0; Mismatches 52; Indels 7; Gaps 2;
QY 535 AGAAAAGCTATTGAAACCGAGATTTCTTAAAAATTCCAAATTCGGAGATTTGCA 594
Db 8 AGAAAACCTATTGAAACCGAGACTTTCTTAAAAATTCCAAATTCGGAGATTTGCA 67

```

```

QY      595 TAACGTGACCAATGATTGGATCAATCCCACTCTAATATATCACTTTCTGAGTTACAGACTT 654
Db      68 TAATGTGACCATGATTATTGGATCAATCCCACTCTAATAGCACTTTTCAGAGTTACAGACTT 127
QY      655 TGAGGAGGAGGAGAGAGATCTTCACTTTCTCTGCCAAACGAAAAAAGGATTTGAACAAAA 714
Db      128 TGAGGAGATGAGTGAAGACCTTTCACTTTCTCTACCAAGTGAAGAAAGGCAATTTGAACAAA 187
QY      715 TGAACAGTGGGTGGTCCCTCAAGTGAAGTGAAGAGACCCGTCACGCCAGACA---AGC 771
Db      188 CGAGCAGTGGGTGGTCCCTCAAGTGAAGTGAAGAGACCCGTCACGCCAGAGCAGC 247
QY      772 AAGTGAGGAGAACTTTCCCAATTAATGACTATATCACTTCAAAATGGAATAGATTTGATCCCAT 831
Db      248 AAGTGAGGAGAACTTTCCCAATTAATGACTATATCACTTCAAAATGGAATAGATTTGATCCCAT 307
QY      832 GCTGATCAGAGAGAGTATTGTTGTTATTCTGCGCTGAGGCAACCCGCTATTTCGCGCG 891
Db      308 GTTGGATGAGAGAGGTTACTGTTGTTATTCTGCGCTGAGGCAACCCGCTACTGTGCGCG 367
QY      892 CGTCTGTGAACCTTTTACTAGGCTACTACCATATCCATCTCTCTACCAAGGAGGAGCT 951
Db      368 CGTCTGTGAACCTTTTACTAGGCTACTACCATATCCATCTCTCTACCAAGGAGGCGGT 427
QY      952 CATCTGTCTGTCTATCATGCTTGTAACTGTGGTGGCGCCGCTGCTCTACCAAGGAGGAGCT 1011
Db      428 TATCTGTCTGTCTATCATGCTTGTAACTGTGGTGGCGCCGCTGCTCTACCAAGGAGGCTT 487
QY      1012 ATAGGAGGTTTGAGCTCAAACTGCTTAACTGCTGG---CAACATATAATAAATGATGTC 1067
Db      488 ATAGCAAGTCTGAGCTCGAGTCTTAACTTCTGGCATCCACATATAACAAATGATGTC 547
QY      1068 TATTCATGAATTTCTGCTCTATGAGGCTCTGCGCCCTGCTAGCCAGCTCTCCAGATTA 1127
Db      548 TATTCATGAATTTCTGCTCTATGAGGCTCTGCGCCCTGCTAGCCAGCTCTCCAGATTA 607
QY      1128 CTT 1130
Db      608 CTT 610

RESULT 30
BI304069/c 636 bp mRNA linear EST 20-JUL-2001
LOCUS UI-R-DR0-cjc-o-14-0-UI.s1 UI-R-DR0 Rattus norvegicus cDNA clone
DEFINITION UI-R-DR0-cjc-o-14-0-UI 3', mRNA sequence.
ACCESSION BI304069
VERSION BI304069.1 GI:14980349
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 636)
Bonaldio, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL 97044477
MEDLINE 889548
PUBMED
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site

```

and the oligo-dT track served to verify it as a clone from the non-normalized osteoblast library cDNA library Preparation: M.B. Soares lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes.

FEATURES

Location/Qualifiers
1..536
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DR0-cjc-o-14-U-UI"
/dev_stage="ADULT"
/lab_host="DHI0B (Life Technologies)"
/clone_lib="UI-R-DR0"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; the UI-R-DR0 library is a non-normalized Rat Osteoblast library constructed in pT7T3D PAC vector according to the procedure described by Bonaldo, Lennon & Soares (Genome Research Genome 6: 791-806, 1996). The oligonucleotide used to prime first strand synthesis contained the sequence tag CCAGG between the Not I cloning site and dT18 stretch.
TAG TISSUE=osteoblast
TAG_LIB=UI-R-DR0
TAG_SEQ=AGATATCA"

ORIGIN

Query Match 40.6%; Score 485.4; DB 12; Length 636;
Best Local Similarity 87.0%; Pred. No. 4.9e-83;
Matches 561; Conservative 0; Mismatches 71; Indels 13; Gaps 2;
553 CCGAGATTCTTAAAAATTCAGAAATTCGAGATTGTGATAACGTCACCATGTATTG 612
636 CAGAGACTTCTGAAAAATTCATAATTCGGAATTTGGCAATGTGACTAIGTACTG 577
613 GATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAGACTTTTGAGGAGGAGGAAGA 672
576 GATCAATCCCACTCTAATAGCAGTTTCAGAAATACAGGACTTTTGAGGAGGATGCTGAAGA 517
673 TCCTCACTTCTTCCGCAACCAAAAAAGGAGTGAACAAATGAACAGTGGGTGGTCCC 732
516 TCCTCACTTCTTCCGCAACCAAAAAAGGAGTGAACAAATGAACAGTGGGTGGTCCC 457
733 TCAAGTGAAGTGAAGAGCCCGTCACGCCAGCAAGCAAGTGAAGAACTTCCAAT 792
456 ACAAGTGAAGTGAAGAGCCCGTCACGCCAGCAAGCAAGTGAAGAACTTCTCTGT 397
793 AAATGACTACTGAAAATGGAATAGAAATTTGATCCCATCTGCTGATGAGAGGTTATTG 852
396 TAATGACTACTGAAAATGGAATGGAATTTGATCCCATCTGCTGATGAGAGGTTACTG 337
853 TTGTAATTTACTGCGCTCGAGCAACCGTATTCGCCGCTGCTGTAACCTTTACTAGG 912
336 TTGTAATTTACTGCGCTCGAGCAACCGTACTCTCCGCGAGGCTCTGTAACCTTTACTAGG 277
913 CTACTACCATATPCCATATCTGTACCAAGAGGACGAGTCACTGCTGCTCATCATGCCC 972
276 CTACTACCATATCCCTACTGCTTACCAAGAGGTCAGATCACTGCTGCTCATCATGCCC 217
973 TTGTAATTTACTGCGCTCGAGCAACCGTATTCGCCGCTGCTGTAACCTTTACTAGG 1032
216 TTGCAACTGGTGGGTGGCCGCAATGCTTGGGAGAGTCTAATAGAGGTTTGAGCTCAAAAT 157
1033 GCTTAACTGCT-----GGCAACATATAATAATGCAATGCTATTCATGAATTTCTGCCTA 1088
156 GCTTAACTTTTGTAGCAACATATAATTAATGCAATGCTATTCATGAATTTCT----- 102
1089 TGAGGCATCTGGCCCTGGTAGCCAGCTCTCCAGAAATTAATTTGTAGGTAATTTCTCTCT 1148
101 -----GCAATTTGCTCCAGTAGCTATCTCCAGAAATTAATTTGTAGGTAATTTCTCTCT 46

QY 1149 CATGTTCTTAATAAACTTCTACATTATCACCAAAAAA 1193
Db |||||

45 CGTGTTCTTAATAAACGCTCTACATCATCAAAAAA 1

RESULT 31

AI123839/c

LOCUS

DEFINITION

q47e02.x1 Soares fetal heart NbHH19W Homo sapiens cDNA clone

IMAGE:1692506 3' similar to SW:CHM1_BOVIN PL7404 CHONDRMODULIN-I

PRECURSOR ; mRNA sequence.

AI123839

AI123839.1 GI:3539605

EST.

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 479)

NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 591 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 440.

Location/Qualifiers

1..479

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:1692506"

/sex="unknown"

/dev_stage="19 weeks"

/lab_host="DHI0B (ampicillin resistant)"

/clone_lib="Soares fetal heart NBHH19W"

/notes="Organ: heart; Vector: pT7T3D (Pharmacia) with a

modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo (dT) primer [5'

TGTTACCAATCTGAAGTGGAGCGGCGCATCTTTTTTTTTTTT 3']

double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified pT7T3 vector

(Pharmacia). Library went through one round of

normalization to a Cot = 5. Library constructed by

M.Patima Bonaldo. This library was constructed from the

same fetus as the fetal lung library, Soares fetal lung

NBHH19W."

ORIGIN

Query Match

Best Local Similarity 99.6%; Score 474.8; DB 9; Length 479;

Matches 476; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 699 AAGGGATTGAACAAATGAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAGACCCGTC 758

Db |||||

479 AAGGGATTGAACAAATGAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAGACCCGTC 420

QY 759 ACGCCAGACAACAGTGAAGAGAACTTCCAAATAAAGACTATATCTGAAAATGGAATAG 818

Db |||||

419 ACGCCAGACAACAGTGAAGAGAACTTCCAAATAAAGACTATATCTGAAAATGGAATAG 360

QY 819 AATTGTATCCCATGCTGGATGAGAGGTTATGTTGTTATTCTGCGCTCGAGGCAACC 878

Db |||||

359 AATTGTATCCCATGCTGGATGAGAGGTTATGTTGTTATTCTGCGCTCGAGGCAACC 300

QY 879 GCTATTGGCGCGCGCTCTGTGAACCTTTACTAGGCTACTTACCCTATATCCATCTGCTACC 938

Db |||||

299 GCTATTGGCGCGCGCTCTGTGAACCTTTACTAGGCTACTTACCCTATATCCATCTGCTACC 240

/lab_host="DH10B (ampicillin resistant)"/
/clone_lib="Soares fetal heart NBH19W"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTTACCAATTCGAAGTGGAGCGCCGATCTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fátima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBH19W."

ORIGIN

Query Match	39.1%;	Score 468;	DB 13;	Length 468;
Best Local Similarity	100.0%;	Pred. No. 1.2e-79;		
Matches 468;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	633	CAGTTTCTGAGTTACAAGACTTTTGAGGAGGAGGAGAAAGTCTTTCAC	692	
Db				
QY	1	CAGTTTCTGAGTTACAAGACTTTTGAGGAGGAGGAGAAAGTCTTTCAC	60	
Db				
QY	693	AAAAAAAAAGGGATTGAACAAAAATGAACAGTGGGTGGTCCCTCAAGTCAAAAGTAGAGAGA	752	
Db				
QY	61	AAAAAAAAAGGGATTGAACAAAAATGAACAGTGGGTGGTCCCTCAAGTCAAAAGTAGAGAGA	120	
Db				
QY	753	CCCGTCACGCCAGACAAAGCAAGTGAGAGAAACTTCCAATAAATGACTATACATGAAAAATG	812	
Db				
QY	121	CCCGTCACGCCAGACAAAGCAAGTGAGAGAAACTTCCAATAAATGACTATACATGAAAAATG	180	
Db				
QY	813	GAAATAGAAATTTGATCCCATGCTCGATGAGAGAGGTTATTGTTGATTTACTGCGCTCGAG	872	
Db				
QY	181	GAAATAGAAATTTGATCCCATGCTCGATGAGAGAGGTTATTGTTGATTTACTGCGCTCGAG	240	
Db				
QY	873	GCAACCGCTATTGCGCCGCGTCTGTGAACTTTTACTAGCTACTACCCATATCCATACT	932	
Db				
QY	241	GCAACCGCTATTGCGCCGCGTCTGTGAACTTTTACTAGCTACTACCCATATCCATACT	300	
Db				
QY	933	GCTACCAAGGAGGACGAGTCACTGTGCTGTGTCATCATGCTTGTACTGGTGGGTGGCCC	992	
Db				
QY	301	GCTACCAAGGAGGACGAGTCACTGTGCTGTGTCATCATGCTTGTACTGGTGGGTGGCCC	360	
Db				
QY	993	GCATGCTGGGAGGGGTCTAATAGGAGGTTTGAGCTCAAAATGCTTAAACTGCTGGCAACAT	1052	
Db				
QY	361	GCATGCTGGGAGGGGTCTAATAGGAGGTTTGAGCTCAAAATGCTTAAACTGCTGGCAACAT	420	
Db				
QY	1053	ATAATAAATGCATGCTATTCAATGAATTTCTGCGTATGAGGCATCTGG	1100	
Db				
QY	421	ATAATAAATGCATGCTATTCAATGAATTTCTGCGTATGAGGCATCTGG	468	
Db				

RESULT 34

BF121576	888 bp	linear	EST 24-OCT-2000
60175639771 NCI_CGAP_Mam5	Mus musculus	cDNA clone	IMAGE:3986395 5',
RNA sequence.			
BF121576			
BF121576.1	GI:10960616		
EST.			
Mus musculus (house mouse)			
Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
1 (bases 1 to 888)			
NIH-MGC http://mgc.nci.nih.gov/			
National Institutes of Health, Mammalian Gene Collection (MGC)			
Unpublished (1999)			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-r@mail.nih.gov			
Issue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys			
cDNA Library Preparation: Life Technologies, Inc.			

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM9191 row: i column: 20
 High quality sequence stop: 677.
 Location/Qualifiers
 1..888
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:3986395"
 /tissue_type="tumor, gross tissue"
 /dev_stage="7 months"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Mam5"
 /note="Organ: Mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Robin Humphreys,
 NTH"

[illegible]

ORIGIN

Query Match	38.7%;	Score 463.6;	DB 10;	Length 888;
Best Local Similarity	83.6%;	Prod. No. 6.5e-79;		
Matches 632;	Conservative	0;	Mismatches 114;	Indels 10; Gaps 9;
QY	53	GAGACCATGGCAAGAATCCTCCAGAGAAATGTGCAAGACTGTACATCTTAAATCGAGAA	112	
Db	1			
QY	113	GCITTTTAAATCCAGAAAAATATGTAAATCACTTAAGATTTGTGCACCTGGTCTTTGGTATC	172	
Db	61			
QY	173	CTGGCCCTAACTCTAAATTTGTCCTGTTTTGGGGAGCAAGCACTTCTGCGCGGAGGTACCC	232	
Db	119	CTGGCCTTAACTCTAAAT - GTCCCTGTTTGGGGAGCAACACACTTCTGCGCCGAGTATCC	176	
QY	233	AAAAAGCCTATGACATGGAGCACATTTCTACGCAATGGAGAGAAGAAGATTTAC	292	
Db	177			
QY	293	ATGGAATTCATCTCTGTACAGAACTGGAATATTTCAGAAACGGAATGCACTGATGAA	352	
Db	237			
QY	353	ACATTTGGAAGTCACGACATTTTAAAAACGGGATACACTGGCACTCTACTTCGTGGGTCTTCAA	412	
Db	296	ACATTTGGAAGTCACGACATTTTAAAAATGGGATACACTGGCATCTACTT - GTAGGTCTTCAA	354	
QY	413	AAATGTTTTATCAAACTCAGATTAAGTGATTCCTGAATTTTCTGAACACAGAAGAGGAA	472	
Db	355			
QY	473	ATAGATGAGAATCAAGAAATATTACCAACTTTCCTTTGAAACAGTCAGTGATTTGGGTCCCA	532	
Db	415			
QY	533	GCAGAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCGAGAGATTGT	592	
Db	474	GCAGAAAGCCTATTGAAAACAGAGACTTCTGAAAAATTTCAAATTTCTGGAGATTG -	532	
QY	593	GATAAGTCGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTCAGTTTACAAGAC	652	
Db	533			
QY	653	TTTGAGGAGGGGAG - AGATCTTTCAC - TTTCTCGCCCAACGAAAAAAAAGGATTGAAC	710	
Db	593			

```

QY 711 AARATGAACAGTGGTGGTCCCTCAAGTGAAGTAGAGAACCCGCTCACCCAGACAG 770
Db 653 AAGATGAGCAGAGGTGGTCCGCAAGTTGAGGGGGGGAACCCGCAACAGAGAGCAG 712
QY 771 CAAGTGAAGGAAGAACTTCCAAATAAATGACTATAGT 806
Db 713 CGAGGGAAGAACTTCATTAAAGACCAATATGGCACTG 748

RESULT 35
BF439260/c
LOCUS
DEFINITION
nab61h07.x1 Soares NSF F8.9W OT PA P S1 Homo sapiens cDNA clone
IMAGE:3270397 3' similar to IR:Q9UJG0 Q9UJG0 D44797.1 ; mRNA
sequence.
ACCESSION
BF439260
VERSION
BF439260.1 GI:11451777
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 462)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco.
FEATURES
Location/Qualifiers
1..462
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3270397"
/lab_host="PH108"
/clone_lib="Soares NSF F8.9W OT PA P S1"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NBHSF pool 1:
309384-310919, 323208-325895 Soares NB2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NB2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NBHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."
ORIGIN
Query Match 38.4%; Score 460; DB 10; Length 462;
Best Local Similarity 100.0%; Pred. No. 4e-78;
Matches 460; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 717 AACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAACCCGCTCACGCCAGACAGCAAGTGT 776
Db 462 AACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAGAACCCGCTCACGCCAGACAGCAAGTGT 403
QY 777 AGGAAGAACTTCCAATAAATGACTATATCTGAAATGAATGAATGATATTCCTCCATGCTGG 836
Db 402 AGGAAGAACTTCCAATAAATGACTATATCTGAAATGAATGAATGATATTCCTCCATGCTGG 343
QY 837 ATGAGAGAGGTATTTGTTGTTATTTACTGCGCTCCAGGCCAACCGCTATTTGCGCGCGCTCT 896
Db 342 ATGAGAGAGGTATTTGTTGTTATTTACTGCGCTCCAGGCCAACCGCTATTTGCGCGCGCTCT 283

```

```

QY 897 GTGAACCTTTTACTAGGCTACTACCCATATCCATATCTGCTACCAAGGAGGAGGTCATCT 956
Db 282 GTGAACCTTTTACTAGGCTACTACCCATATCCATATCTGCTACCAAGGAGGAGGTCATCT 223
QY 957 GTCGTGTCATCATGCTTGTAACTGTTGGTGGCCCGCATGCTGGGAGGGGTCTAATAGG 1016
Db 222 GTCGTGTCATCATGCTTGTAACTGTTGGTGGCCCGCATGCTGGGAGGGGTCTAATAGG 163
QY 1017 AGGTTTGAGCTCAAAATGCTTAAACTGCTGGCAACATATAATAATGCTATGCTATTCATG 1076
Db 162 AGGTTTGAGCTCAAAATGCTTAAACTGCTGGCAACATATAATAATGCTATTCATTCATG 103
QY 1077 AATTTCTCCCTATGAGGCATCTGGCCCTGCTAGCCAGCTCTCCAGAAATTTACTTCTAGT 1136
Db 102 AATTTCTCCCTATGAGGCATCTGGCCCTGCTAGCCAGCTCTCCAGAAATTTACTTCTAGT 43
QY 1137 AATTCCTCTCTTCAATGTTCTAATAAATCTTACATTTATCA 1176
Db 42 AATTCCTCTCTTCAATGTTCTAATAAATCTTACATTTATCA 3

T12179 465 bp mRNA linear EST 28-NOV-1994
A533F Heart Homo sapiens cDNA clone A533 similar to Chondromodulin,
mRNA sequence.
ACCESSION
T12179
VERSION
T12179.1 GI:596883
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 465)
Liew, C.C., Hwang, D.M., Fung, Y.W., Laurensen, C., Cukerman, E.,
Tsui, S. and Lee, C.Y.
A catalogue of genes in the cardiovascular system as identified by
expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 91, 10645-10649 (1994)
JOURNAL
MEDLINE
95024171
PUBMED
7938007
COMMENT
Other ESTs: A533R
Contact: Liew CC
Brigham and Women's Hospital
Harvard Medical School
75 Francis St. Boston, MA 02115, USA
Tel: 6177328915
Fax: 6177950995
Email: cliw@rics.bwh.harvard.edu
Similar to bovine chondromodulin
Seq primer: GGTGGCAGCACTCTCTGGAGCC.
FEATURES
Location/Qualifiers
1..465
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="A533"
/lab_host="E.coli Y1090"
/clone_lib="Heart"
/note="Vector: Lambda gt11; Site_1: EcoRI; Site_2: EcoRI"
ORIGIN
Query Match 36.4%; Score 435.8; DB 14; Length 465;
Best Local Similarity 99.1%; Pred. No. 1.7e-73;
Matches 459; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 451 ATTTCTTGACCAAGAGGAAATAGATGAATGAAGAAATACCAACAACTTTCTTTGA 510
Db 461 AATTTCTGAAACCAAGAGGAAATAGATGAATGAAGAAATACCAACAACTTTCTTGA 403
QY 511 ACAGTCAGTGTATTTGGTCCAGCAAAAAGCCCTATTGAAACCGAGATTTCTTAAAAA 570

```



```
Db 402 ACAGTCAGTGATTTGGGTCCCGACGAGAAAAGCCTATTGAAAAACCGAGATTTTCTTTAAAAA 343
QY 571 TTCAAAAATTCGGAGATTTGTGATAAAGCTGACCATGATTGATCAATCCACACTCTAAT 630
Db 342 TTCAAAAATTCGGAGATTTGTGATAAAGCTGACCATGATTGATCAATCCACACTCTAAT 283
QY 631 ATCAGTTTCTGAGTTTACAAGACTTTGAGAGAGGAGGAGAAAGATCTTCACTTTCTTCGCCAA 690
Db 282 ATCAGTTTCTGAGTTTACAAGACTTTGAGAGAGGAGGAGAAAGATCTTCACTTTCTTCGCCAA 223
QY 691 CGAAAAAAGGAGTTGAACAAAATGACAGTGGGTGGTCCCTCAAGTCAAAAGTAGAGAA 750
Db 222 CGAAAAAAGGAGTTGAACAAAATGACAGTGGGTGGTCCCTCAAGTCAAAAGTAGAGAA 163
QY 751 GACCGGTCAACCGACGACGAGCAAGTGAAGAACTTCCAATTAATGACTACTAGAAA 810
Db 162 GACCGGTCAACCGACGACGAGCAAGTGAAGAACTTCCAATTAATGACTACTAGAAA 103
QY 811 TGGATAGAAATTTGATCCATGCTGGATGAGAGAGGTTATTGTTGATTTACTGCCGTCG 870
Db 102 TGGATAGAAATTTGAT-CCATGCTGGATGAGAGAGGTTATTGTTGATTTACTGCCGTCG 44
QY 871 AGGCAACCGCTATTGCCCGCGCTGTGACCTTTACTAGGC 913
Db 43 AGGCAACCGCTATTGCCCGCGCTGTGAACTTTACTAGGC 1

RESULT 37
BF123957
LOCUS 601759641F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4022674 5',
DEFINITION mRNA sequence.
ACCESSION BF123957
VERSION BF123957.1 GI:10962997
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9280 row: a column: 11
High quality sequence stop: 703.
FEATURES
Location/Qualifiers
1..856
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4022674"
/tissue_type="tumors, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam5"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

35.5%; Score 424.6; DB 10; Length 856;

ORIGIN
Query Match
```

```
Best Local Similarity 89.5%; Pred. No. 1.9e-71;
Matches 513; Conservative 0; Mismatches 54; Indels 6; Gaps 5;

QY 11 CTCTCAGTCTCTCTAAAGCAAGAAAGTAGTACTGTGCTCGAGAGACCATGCAAGAAGAT 70
Db 142 CTCTCAGTCTCTCTAAAGCAAGAAAGTAGTACTGTGCTCGAGAGACCATGCAAGAAGAT 201
QY 71 CCTCCAGAGANTTGTGAAGACTGTGCACATTCTAATGAGAGCTTTTAAATCCAGAAA 130
Db 202 CCTCCAGAGANTTGTGAAGACTGTGCACATTCTAATGAGAGCTTTTAAATCCAGAAA 261
QY 131 ATATGTTAATCACTTAAGATTGTGGACTGTGTGTTGTTATCTCTGCGCTTAACTTAA 190
Db 262 ATATGTTAATCACTGAAGATTGTGGACTGTG-TTGGTATCTCTGCGCTTAACTTAA 320
QY 191 GTCTGTTTGGGGGAGCAAGCACTTCTGGCCGAGGTACCCAAAAAGCCTATGACATG 250
Db 321 GTCTGTTTGGGGGAGCAAGCACTTCTGGCCGAGGTATCCAAGAAAACCTATGACATG 380
QY 251 GAGCACACTTTCTACAGCAATGGAGAGAGAAAGATTTACATGGAATTTGATCTCTGT 310
Db 381 GAGCACACTTTCTACAGCAATGGAGAGAGAAAGATTTTACATGGAATTTGATCTCTGT 440
QY 311 ACCAGAACTGAATATTTCAGAAAGCGGAAATGCGCACTGATGAAACATTTGGAAGTGCACGAC 370
Db 441 ACCAGAACTGAATATTTCAGAAAGTGGAAATGGCACTGATGAAACA-TGGAAAGTCCATGAC 499
QY 371 TTTAAAAACGGATACATGCGCATCTACTCTGTGGGTCTTCAAAAATGTTTATCAAAACT 430
Db 500 TTTAAAAATGGATACATGCGCATCTACTCTGTGTAGTCTTCAAAAATGCTTTATTAAACT 559
QY 431 CAGATTAAAGTGATTTCTGTAATTTCTGAACCAAGAGAAATAGATGAGAAATGAAGAA 490
Db 560 CAAATCAAAGTGATTTCTGTAATTTCTGAACCAAGAGAAATAGATGAGAAATGAAGAA 619
QY 491 ATTACCAACAACCTTCTTTGAACAGTCACTGATTTGGGTCCCGACAGAAAAGCCTATTGAA 550
Db 620 ATTACTACAACTTTC-TTGAACAGTCACTGTA-TTGGGTCCCGCAGAAAAGCCTAT--GA 675
QY 551 AACCGAGATTTTCTTAAAAATTTCCAAAATTTCTG 583
Db 676 AACAGAGACTTCTGAAAAATCTAAAAATCTGCAG 708

RESULT 38
BF123957
LOCUS 601759641F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4022674 5',
DEFINITION mRNA sequence.
ACCESSION BF123957
VERSION BF123957.1 GI:28841945
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Dr. Leslie L. Heckert
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM41 row: m column: 14
High quality sequence stop: 422.
FEATURES
Location/Qualifiers
1..741
source
```

```
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30281869"
/tissue_type="primary cultures of sertoli cells"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH_MGC_165"
/notes="Organ: testis; Vector: pDNR-LIB; Site 1: Sfil
(ggcatatggcc); Site 2: Sfil (ggcgcctcgcc); 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGGCGCATATGACC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCGGCACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4 kb
(range 0.6-3.5 kb). 15/15 colonies contained inserts by
PCR. This library was enriched for full-length clones and
was constructed by Clontech Laboratories (Palo Alto, CA).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 35.0%; Score 419; DB 14; Length 741;
Best Local Similarity 88.7%; Pred. No. 2.4e-70;
Matches 489; Conservative 0; Mismatches 55; Indels 7; Gaps 3;

QY 11 CTCTCAGTCTCTCAAGCAAGAAAGACTGTGTGCTGAGAGACCATTGGCAAGAAT 70
DB 11 |||||
DB 40 CTCTCAGTCTCTCAAGCAGGAAAGACCGGTGTGCTGGAGACCATTGGCAAGAAT 99
DB 40 |||||
QY 71 CTCCAGAGAATGTGAAGCTGTGCATCTTAATTCGAGAAGCTTTAAATCCAGAAA 130
DB 71 |||||
DB 100 CTCCAGAGAATGTGAGGGCTGTGCATCTTAATTCGAGAAGCTTTAAATCCAGAAA 159
DB 100 |||||
QY 131 ATATGTAATCACTTAAGATTGTGACGTGTGTTGGTATCTCGCCCTTAATCTTAAT 190
DB 131 |||||
DB 160 ATATGTAATCACTGAAGATTGTGACGTGTGTTGGTATCTCGCCCTTAATCTTAAT 219
DB 160 |||||
QY 191 GTCTCTGTTTGGGGAGCAAGCACTTCTGCGCGGAGGTACCCAAAAGCCTATGACATG 250
DB 191 |||||
DB 220 GTCTCTGTTTGGGGAGCAAGCACTTCTGCGCGGAGGTATCCAAAGAAACCTATGACATG 279
DB 220 |||||
QY 251 GAGCACTTTTACAGCAATGAGAGAGAAGAACTTACATGGAATTTGATCTCTGTG 310
DB 251 |||||
DB 280 GAGCACTTTTACAGCAATGAGAGAGAAGAACTTACATGGAATTTGATCTCTGTG 339
DB 280 |||||
QY 311 ACCAGAACTGAATATTTACAGAGCGAAATGGCACTGATGAAACATTTGGAAGTGCACGAC 370
DB 311 |||||
DB 340 ACCAGAACTGAATATTTACAGAGCGAAATGGCACTGATGAAACATTTGGAAGTGCACGAC 399
DB 340 |||||
QY 371 TTTTAAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTTTATCAAAACT 430
DB 371 |||||
DB 400 TTTTAAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTTTATCAAAACT 459
DB 400 |||||
QY 431 CAGATTAAAGTGAATCTCTG-AAATTTCTGAACAGAGAGGAATAGATCAGAAATGAAGA 489
DB 431 |||||
DB 460 CAAATCAAAGTGAATCTCTGAAATGTTCTGAACAGATGAAGAAATAGATGAGAAATGAAGA 519
DB 460 |||||
QY 490 AATTAC--ACAACTTCTTTGAACAGTCAAGTGAATTT---GGGTCCAGCAGAAAAAGCC 543
DB 490 |||||
DB 520 AAATTACTTACAATTTCTTTGACAGTCAAGTGAATTTGGGGTTCCGGCCAAAAAGCC 579
DB 520 |||||
QY 544 TATTGAAAACC 554
DB 544 |||||
QY 580 TATTGGAAC 590
DB 580 |||||

RESULT 39
BX638501/c
LOCUS
DEFINITION BX638501 pBluescript Lion Mus musculus cDNA clone LIONp462G08420
3', mRNA sequence.
ACCESSION BX638501
VERSION BX638501.1 GI:33618376
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
```

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 521)
Henrich, J., Hermanns, J., Kranz, H., Loebbert, R., Schluter, T.,
Schuette, D., Weindel, M., Heil, O., Ebert, L., Neubert, P., Peters, M.,
Radelof, U., Schneider, D. and Korn, B.
Mouse ArrayTAG cDNA (LION)
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; LIONp462G08420.
RZPDLIB;
Mouse ArrayTAG cDNA (LION)
http://www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=4
62 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
RP: CAGAAACAGCTATGAC.
Location/Qualifiers
1..521
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="LIONp462G08420"
/lab_host="DH10B"
/clone_lib="pBluescript Lion"

ORIGIN
Query Match 34.2%; Score 409.4; DB 13; Length 521;
Best Local Similarity 87.9%; Pred. No. 1.9e-68;
Matches 457; Conservative 0; Mismatches 59; Indels 4; Gaps 1;

QY 668 GAAGATCTTCACCTTCTGCTCCCAACGAAAAAAGGATTGAAACAAATGACAGTGGTG 727
DB 668 |||||
QY 728 GTCCCTCAAGTAGAAGTAGAAGACCCGTACAGCCAGCAACGCAAGTAGAGGAAGACTT 787
DB 728 |||||
QY 460 GTCCCGCAAGTGAAGGTGGAGAGACCCGCCACACAGCAAGCAGGAGGAGACCTT 401
DB 460 |||||
QY 788 CCAATAAAGACTATATCTGAAATGGAATGAAATTTGATCCCATGCTGGATGAGAGAGT 847
DB 788 |||||
QY 400 CCTATAAATGACTATCTGAAATGGAATTTGATCCCATGCTGGATGAGAGAGT 341
DB 400 |||||
QY 848 TATTGTTGATTTTACTGCGGTGAGGCAACCGCTATTGCGCGCGGTCTGTGAACCTTTA 907
DB 848 |||||
QY 340 TACTGTTGATTTTACTGCTGAGGCAACCGTTACTGCGCGCGGTCTGTGAACCTTTA 281
DB 340 |||||
QY 908 CTAGGCTACTACCCATATCCATCTACTGCTACAGGAGGAGAGTCACTCTGCTGTCATC 967
DB 908 |||||
QY 280 CTAGGCTACTACCCATATCCCTTACTGCTACCAAGGAGGTGAGTCACTCTGCTGTCATC 221
DB 280 |||||
QY 968 ATGCTTTGTAACCTGTTGGGTGGCCCGCATGCTGGGAGGGGTCTTAATAGGAGGTTTGAGT 1027
DB 968 |||||
QY 220 ATGCTTTGCAACTGTTGGGTGGCCCGCATGCTTTGGGAGAGTCTAATAGGAAGATTGAGT 161
DB 220 |||||
QY 1028 CAAATGCTTAAACTGCT---GGCAACATATATAATGATGATGCTATTCATGATTTCT 1083
DB 1028 |||||
QY 160 CCNCGCTTAACTTCTGTTAGCCATATATATAATGATGCTATTCCTCCATGAATTTCT 101
DB 160 |||||
QY 1084 GCCTATGAGCACTTGGCCCTGTTAGCCAGCTCTCCAGAAATTAATTTGTTAGTAAATTCCT 1143
DB 1084 |||||
QY 100 GCCTATGAGCAATTTGCTCCAGTAGCCCTATTCCTTCAGAAATTAATTTGTTAGTAAATTCCT 41
DB 100 |||||
QY 1144 CTCTTCAATGTTTAATAAACTTTCTACATTTATCACCAAAAA 1183
DB 1144 |||||


```

Db      40  CTCCTCATGTTCTAATAAACTTCTACATCATCAAAAA 1
RESULT 40
AI039039/c
LOCUS
DEFINITION
  ox30b10.sl Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
  IMAGE:1657819 3', similar to SW:CHM1_BOVIN P17404 CHONDROMODULIN-I
  PRECURSOR i, mRNA sequence.
ACCESSION
AI039039
VERSION
AI039039.1 GI:3278233
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 399)
  NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  This clone is available royalty-free through LLNL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Seq primer: -40ml3 fwd. ET from Amersham
  High quality sequence stop: 389.
FEATURES
  source
  1..399
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:1657819"
  /dev_stage="8-9 weeks"
  /lab_host="DH10B"
  /clone_lib="Soares_total_fetus_Nb2HF8_9w"
  /note="Vector: pT73D-Pac (Pharmacia) with a modified
  polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
  was prepared from mRNA obtained from pooled 8-9 week
  (total) fetus material with a Not I - oligo(dT) primer [5'-
  TGTACCAACTGTAAGTGGGCGCGCTTAATTTTTTTTTTTTTTTT 3'].
  Double-stranded cDNA was ligated to Eco RI adaptors
  (Pharmacia), digested with Not I and cloned into the Not I
  and Eco RI sites of the modified pT73 vector. Library
  went through one round of normalization, and was
  constructed by Bento Soares and M. Fatima Bonaldo. "
ORIGIN
  Query Match 32.2%; Score 385; DB 9; Length 399;
  Best Local Similarity 99.7%; Pred. No. 9.5e-64;
  Matches 396; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY      780  AAGAAGCTTCCAAATAAATGACTATACTGAAATGGAATGGAATTTGATCCCATGCTGGATG 839
Db      399  AAGAACTTCCAAATAAATGACTATACTGAAATGGAATGGAATTTGATCCCATGCTGGATG 340
QY      840  AGAGAGTTATTGTTGTTATTACTGCGTCGAGGCAACCGCTATTGCGCGCGCTGCTGTG 899
Db      339  AGAGAGTTATTGTTGTTATTACTGCGTCGAGGCAACCGCTATTGCGCGCGCTGCTGTG 280
QY      900  AACCTTTACTAGCTACTACCATATCCATCTGCTACCAAGGAGGACGAGTCATCTGTC 959
Db      279  AACCTTTACTAGCTACTACCATATCCATCTGCTACCAAGGAGGACGAGTCATCTGTC 220
QY      960  GTGTCATCATGCTTGTAACTGTTGGTGGCCCGCATGCTGGGAGGCTTAATAGAGG 1019
Db      219  GTGTCATCATGCTTGTAACTGTTGGTGGCCCGCATGCTGGGAGGCTTAATAGAGG 160
QY      1020  TTTGAGCTCAATGCTTAACTGCTGGCAACATATAATAATGATCATGCTATTCAATGAAT 1079
Db      159  TTTGAGCTCAATGCTTAACTGCTGGCAACATATAATAATGATCATGCTATTCAATGAAT 100
QY      1080  TTCTGCCTATGAGGCATCTGGCCCTTGGTAGGCAGCTCTCCAGAATTACTTGTAGGTAAT 1139

```

```

Db      99  TTCT-CCTATGAGGCATCTGGCCCTGGTAGCAGCTCTCCAGATTACTTGTAGTAAT 41
QY      1140  TCCTCTCTTCATGTTCTTAATAAACTTCTACATTATCA 1176
Db      40  TCCTCTCTTCATGTTCTTAATAAACTTCTACATTATCA 4

RESULT 41
CB202990
LOCUS
DEFINITION
  CB202990 11290775 NIH_MGC_135 Mus musculus cDNA clone
  IMAGE:30139809 5', mRNA sequence.
ACCESSION
CB202990
VERSION
CB202990.1 GI:28238967
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Mus musculus
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 1224)
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  Tissue Procurement: Dr. David Rowe
  cDNA Library Preparation: Invitrogen Corp
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Agencourt Bioscience Corporation
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: NDAM0044 row: n column: 10
  High quality sequence start: 86
  High quality sequence stop: 334.
FEATURES
  source
  1..1224
  /organism="Mus musculus"
  /mol_type="mRNA"
  /db_xref="taxon:10090"
  /clone="IMAGE:30139809"
  /lab_host="DH10B (phage-resistant)"
  /clone_lib="NIH_MGC_135"
  /note="Vector: pCMVSPORT6.1; Site 1: EcoRV; Site 2: NotI;
  Normalized full-length enriched library from pooled mouse
  embryonic limb, maxilla and mandible, day 12.5 13.5,
  14.5, and 15.5 (size selected for the 0.5-1 kb fragments)
  Cloned directionally, priming method: Oligo-dT. cDNA
  enrichment: >1k bp, Average insert size 1.6k bp.
  Normalization (Cot value): 7.5 kb. Priming sequence:
  5'GACTAGTTCTAGATCGGCGGCCCTT3' Tissue contributed
  by, David Rowe. Library constructed by ResGen, Invitrogen
  Corp."
ORIGIN
  Query Match 31.7%; Score 379.8; DB 14; Length 1224;
  Best Local Similarity 85.0%; Pred. No. 6.5e-63;
  Matches 472; Conservative 0; Mismatches 77; Indels 6; Gaps 4;

QY      12  TCTCAGTCTCTCAAGCA-AGGAAGAGTACTGTGCTGAGAGACCATGGCAAGAAT 70
Db      64  TCTCCGGGATCTCAAGCAGGGGAAAGAGCATCTGTGCTGGAGACCATGGCAAGAAT 123
QY      71  CCTCCAGAGAAATGTGAAGACTGTACATTCTAAATGCAGAGCTTTTAAATCCAGAAA 130
Db      124  CCTCCAGAGAACTGTGAGGGCTGTACATCTAAATGCAGAGCTCTGAATCTAAGAAG 183
QY      131  ATATGTAATCACTTAAGATTTTGGACTGGTGTGGTATCTGCGCCCTAACTCTAATT 190
Db      184  ATATGTAATCACTGAAGATTTTGGACTAGTGTGGTATCTGCGCCCTAACTCTAATT 243
QY      191  GTCTGCTTTTGGGGAGCAGCACTTCTGGCGGAGGTACCCAAAAAGCCCTATGACATG 250

```

Best Local Similarity 83.3%; Pred. No. 1.3e-62;		Matches 539; Conservative 0; Mismatches 94; Indels 14; Gaps 9;	
Db	244		
Qy	251	GAGCAGCATTCTTACAGCAATGAGAGAGAGAGAGATTTACATGGAAATTTGATCTCTGTG	310
Db	304	GAGCAGCATTCTTACTGCTACGGCGAGAGAGAGAGATTTACATGGAAATTTGATCTCCATA	363
Qy	311	ACCAGAACTGAAATATTCAGAGCGGAAATGGCACTGATGAAACATTTGGAAGTGCACGAC	370
Db	364	ACCAGAACAGAAATATTCAGAGCTGGAATGGCACTGATGAAACATTTGGAAGTCCATGAC	423
Qy	371	TTTAAAAACGATACACTGCGATCTCTGCTGGTCTTCAAAAATG-TTTTATCAAAAC	429
Db	424	TTTAAAAATGATACACTGCGATCTCTGCTGGTCTTCAAAAATGCTCTTATTAATAAC	483
Qy	430	TCAGATTAAGTGAATCTCTGAAATTTCTGAAACAGAGAGAGAAA-TAGATGAGAAATGAAG	488
Db	484	TCCAATCAAGTGAATCTCTGAAATTTCTGAAACAGAGAGAGAAA-TAGATGAGAAATGAAG	543
Qy	489	AAATTTACACAACTTTCTTT---GAACAGTCAGTGAATTTGGGTCCTCCAGAGAGAAAGCCTA	545
Db	544	AATTTACCAACACTTTTCTTTTTCGAAACAGTCTGAGCATTTGGGTTCTCGCAGAAACCTTA	603
Qy	546	TTGAAAAACCGAGATT 560	
Db	604	TTTGCAAAACCGAGAT 618	
RESULT 42			
BE375361			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
Tissue Procurement: Gilbert Smith, Ph.D.			
cDNA Library preparation: Life Technologies, Inc.			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: LLAM8768 row: i column: 12			
High quality sequence stop: 581.			
Location/Qualifiers			
1..733			
/organism="Mus musculus"			
/mol_type="mRNA"			
/strain="FVB/N"			
/db_xref="taxon:10090"			
/clone="IMAGE:3594323"			
/tissue_type="tumor, biopsy sample"			
/dev_stage="10 months, virgin"			
/lab_host="DH10B"			
/clone_1lb="NCI_CGAP_Mam1"			
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sall;			
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.			
Library constructed by Life Technologies. Investigator			
providing samples: Gilbert Smith, NIH"			
ORIGIN			

FEATURES
source

Location/Qualifiers
1. .460
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1707067"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal heart NbHL19W"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCGAATGGAGCGCGCATCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."

ORIGIN

Query Match 31.4%; Score 376.4; DB 9; Length 460;
Best Local Similarity 99.7%; Pred. No. 4e-62;
Matches 377; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 801 ATACTGAAATGGATAGATGATTTGATCCCATGCTGGATGAGAGAGTTATTGTGTATT 860
Db AGACTGAAATGGATAGATGATTTGATCCCATGCTGGATGAGAGAGTTATTGTGTATT 319

QY 861 ACTGCGTCGAGGCAACCGCTATTGCCCGCGCTGTGTGAACCTTTACTAGGCTACTACC 920
Db ACTGCGTCGAGGCAACCGCTATTGCCCGCGCTGTGTGAACCTTTACTAGGCTACTACC 259

QY 921 CATATCCATCTGCTACCAAGAGAGAGTGATCATCTGTCGTATCATATGCTTTGTAAT 980
Db CATATCCATCTGCTACCAAGAGAGAGTGATCATCTGTCGTATCATATGCTTTGTAAT 199

QY 981 GGTGGTGGCCCGCATCTGGGAGGGTCTAATAGGAGGTTTGAAGCTCAATGCTTAAAC 1040
Db GGTGGTGGCCCGCATCTGGGAGGGTCTAATAGGAGGTTTGAAGCTCAATGCTTAAAC 139

QY 1041 TGCTGGCAACATATAATAATGATGCTGTTTCAATGAATTTCTGCCTATAGGAGTCTGG 1100
Db TGCTGGCAACATATAATAATGATGCTGTTTCAATGAATTTCTGCCTATAGGAGTCTGG 79

QY 1101 CCCCTGTAGCCAGCTCTCCGAATTAATCTGTAGGTAATTCCTCTTCTCAATTAATA 1160
Db CCCCTGTAGCCAGCTCTCCGAATTAATCTGTAGGTAATTCCTCTTCTCAATTAATA 19

QY 1161 AACTTCTACATTATCACC 1178
Db AACTTCTACATTATCACC 1

RESULT 44

AV593196/c
LOCUS AV593196 Bos taurus cartilage fetus Bos taurus linear EST 27-NOV-2001
DEFINITION EICA006A12 3', mRNA sequence.

ACCESSION

AV593196
VERSION AV593196.1
KEYWORDS EST.

SOURCE

Bos taurus (cow)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
1 (bases 1 to 468)
Takaue, A., Hirotsune, S., Itoh, R., Jitohzono, A., Suzuki, H., Aso, H. and Sugimoto, Y.

TITLE

Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs
Nucleic Acids Res. 29 (22), E108 (2001)
21570554
PUBMED 11713328
COMMENT
Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@cocoa.ocn.ne.jp

Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.

FEATURES

source

Location/Qualifiers
1. .468
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="EICA006A12"
/tissue_type="cartilage"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_lib="Bos taurus cartilage fetus"
/note="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A was deleted from a NotI site"

ORIGIN

Query Match 31.0%; Score 371.6; DB 9; Length 468;
Best Local Similarity 90.2%; Pred. No. 3.3e-61;
Matches 422; Conservative 0; Mismatches 39; Indels 7; Gaps 2;

QY 693 AAAAAAAGGATTGAACAAATGAACAGTGGTGGTCCCTCAAGTGAAGTAGAGAAGA 752
Db AAAAAAAGGATTGAACAAACGAGCAGTGGTGGTCCCTCAAGTGAAGTAGAGAAGA 409

QY 753 CCCCTCAGCCAGACA--AGCAAGTGAAGAGAACTTCCAATAAATGACTATATCTGAAA 809
Db CCCCTCAGCCAGACAAGCAAGCAGCAAGTGAAGAGAACTTCCAATAAATGACTATATCTGAAA 349

QY 810 ATGGAATAGAAATTTGATCCCATGCTGATGAGAGAGTTATGTTGTTATTACTGCGGTC 869
Db ATGGAATAGAAATTTGATCCCATGCTGATGAGAGAGTTATGTTGTTATTACTGCGGTC 289

QY 870 GAGCAACCGCTATTGCCGCGGCTCTGTGAACCTTTACTAGGCTACTACCCATATCCAT 929
Db GAGCAACCGCTACTGTGCGCGGCTCTGTGAACCTTTACTAGGCTACTACCCGATCCAT 229

QY 930 ACTGCTACCAAGGAGGACGAGTCACTGTGTCATCATGCGCTTGAAGTGGGTGG 989
Db ACTGCTACCAAGGAGGCGGGTTATCTGTGTCATCATGCGCTTGAAGTGGGTGG 169

QY 990 CCGCATGCTGGGAGGGTCTTAATAGGAGGTTTGAAGCTCAAAATGCTTAAACTGCTGG--- 1046
Db CCGCATGCTAGGAGGGTTTAAAGCACTCTGAGCTCGAGTCTTAAACTTCTGGCAT 109

QY 1047 -CAACATATAAATGCAATGCTATTCAATGAATTTCTGCCTATGAGGCACTGGCCCT 1105
Db CCAACATATAAATGCAATGCTATTCAATGAATTTCTGCCTATGAGGCACTTGGCTCT 49

QY 1106 GGTAGCCAGCTCTCCGAATTAATCTGTAGTAATTCCTCTCTTCATGT 1153
Db GGTAGCCAACTACTCCGGAATTTGTTAGTAATTTCTCTCTTCATGT 1

RESULT 45

CF112025
LOCUS CF112025
DEFINITION
Rattus norvegicus cDNA clone CA2448 5', mRNA sequence.
ACCESSION
CF112025
VERSION
CF112025.1 GI:33169492

KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 433)
Shultz M.A., Zhang, L., Gu, Y.-Z., Baker, G.L., Fannuchi, M.V.,
Padua, A.M., Gurske, W.A., Morin, D., Penn, S.G., Jovanovich, S.B.,
Plopper, C.G. and Buckpitt, A.R.
Gene Expression Analysis in Response to Lung Toxicants: I.
Sequencing and Microarray Development
Unpublished (2003)
Contact: Shultz MA
Dept. of Molecular Biosciences, School of Veterinary Medicine
University of California, Davis
1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA
Tel: 530 752 0793
Fax: 530 752 4698
Email: mashultz@ucdavis.edu
Average Phred score is 20 or better. All poor quality data (Phred <
20) and vector/linker sequence has been removed.
High quality sequence stop: 433.
Location/Qualifiers
1. .433
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="CA2448"
/sex="male"
/tissue type="airway or parenchyma"
/dev_stage="adult"
/clone_lib="Rat lung airway and parenchyma cDNA libraries"
/note="Organ: lung; Vector: pGEM-11zf(-); Site_1: Eco RI;
Site 2: Not I; mRNA was isolated from microdissected rat
lung airways and parenchyma tissues."
FEATURES
source
Query Match 30.8%; Score 369; DB 14; Length 433;
Best Local Similarity 90.8%; Pred. No. 1.1e-60;
Matches 393; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 297 AAATTGATCTGTGACACAGAACTGAATATTTCAGAACGGAAATGGCACTGATGAACAT 356
DB 1 AAATTGATCCATAACACAGACAGAAATATTCAGAACTGGAAATGGCACCGATGAACAT 60
QY 357 TGAAGTGCACGACTTTTAAACACGATACACTGGCATCTACTTCGTGGTCTTCAAAAT 416
DB 61 TGAAGTCCATGACTTTTAAACACGATACACTGGCATCTACTTTGTAGTCTTCAAAAT 120
QY 417 GTTTTATCAAACTCAGATTAAAGTGATTCCTGAATTTCTGACACGAGAGGAATAG 476
DB 121 GCTTTATTAACCTCAATCAAGTAGTAACTCTGAAATTTCTGAAACGAGAGGAATAG 180
QY 477 ATGAGATGAAGAAATACACAACTTTCTTTGAACAGTCAGTCAGTTGGTCCAGCAG 536
DB 181 ATGAGATGAAGAAATTAACACGTTCTTTGAACAGTCAGTCAGTTGGTCCAGCAG 240
QY 537 AAAAGCCTATTGAAACCGAGATTTCTTAAAAATTCGAAATTTCTGGAGATTTGTGATA 596
DB 241 AAAAGCCTATTGAAACAGAGACTTCTCTGAAATTTCTGAAATTTCTGGAGATTTGCGACA 300
QY 597 ACGTGACCATGATTGGATCAATCCCACTCTAATATCAGTTCTTGAGTTTCAAGACTTTG 656
DB 301 ATGTGACTATGTACTGGATCAATCCCACTCTAATATCAGTTCTTGAGTTTCAAGACTTTG 360
QY 657 AGGAGGAGGAGAGATCTTCACTTTCTGCGCAACGAAAAAAGGGATTGAAACAAATG 716
DB 361 AGGAGGATGGTGAGATCTTCACTTTCTTACCACGAAAAAAGGGATTGACCAAGATG 420
QY 717 AACAGTGGGTGGT 729
| | | | |

Db 421 AGCAATGGTGGT 433
RESULT 46
LOCUS A1112003/c
DEFINITION UI-R-Y0-mp-b-11-0-UI.s1 UI-R-Y0 Rattus norvegicus cDNA clone
A1112003 487 bp mRNA linear EST 11-FEB-1999
VERSION UI-R-Y0-mp-b-11-0-UI 3', mRNA sequence.
A1112003.1 GI:3511952
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 487)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence tag present in the cDNA between the NotI site and the
oligo-dr track served to identify it as a clone from the normalized
adult Eye library. cDNA Library Preparation: M. Fatima Bonaldo,
Ph.D. Clone distribution: clones will be available through Research
Genetics
Seq primer: M13 Forward.
Location/Qualifiers
1. .487
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-Y0-mp-b-11-0-UI"
/dev_stage="adult"
/lab host="DH10B (Life Technologies)"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-Y0
library is a subtracted library derived from an
individually-tagged normalized whole-eye (minus the lens)
library. The driver for the subtraction consisted of a
pool of all previous libraries (UI-R-A0, UI-R-A1, UI-R-E0,
UI-R-E1, UI-R-C0, and UI-R-C1). The tag is a string of
3-5 nucleotides present between the Not I site and the
oligo-dr track which allows identification of the library
of origin of a clone within the mixture. The subtracted
library (UI-R-Y0) was constructed as follows: PCR
amplified cDNA inserts from previous library clones from a
which 3' ESTs had been derived were used as a driver in a
hybridization with the normalized whole-eye library in
the form of single-stranded circles. The remaining
single-stranded circles (subtracted library) was purified
by hydroxyapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (Life Technologies) to generate the UI-R-Y0
library. This procedure has been previously described
(Bonaldo, Lennon and Soares, Genome Research 6: 791-806,
1996)"
ORIGIN
Query Match 30.6%; Score 366.8; DB 9; Length 487;
Best Local Similarity 86.9%; Pred. No. 2.7e-60;
Matches 431; Conservative 0; Mismatches 52; Indels 13; Gaps 2;

```

QY 704 ATTGAACAAATGAACTAGTGGTGGTCCCTCAAGTGAAGTAGAGAGACCCGTCAGGCC 763
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 487 ATTGACAGAAATGAGCAATGGTGGTCCCAAGTGAAGTGAAGAGACCCGCGCAC 428
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 764 AGACAGCAAGTGAAGAGAACTTCCAATAAATGACTATCTACTGAAATGAAATGAAATTT 823
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 427 AGACAGCAAGCGAGAGACCTTCCTGTTAATGACTATCTACTGAAATGAAATGAAATTT 368
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 824 GATCCCATGCTGGATGAGAGAGTTATTTGTTATTTACTGCGGTGAGGCAACCGCTAT 883
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 367 GATCCCATGCTGGATGAGAGAGTTACTGTTGTTATTTACTGCGGTGAGGCAACCGCTAC 308
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 884 TGCGCGCGGTCTGTGAACCTTTACTAGGCTACTACCCATATCAATCTACTGTTACCAAGGA 943
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 307 TGCGCGAGGCTGTGAACCTTTACTAGGCTACTACCCATATCAATCTACTGTTACCAAGGA 248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 944 GGAAGAGTCACTGCTGCTGCTATCATGCTCTGTAAGTGGTGGTGGCGGATGCTGGGG 1003
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 247 GGTGAGTCACTGCTGCTGCTATCATGCTCTGTAAGTGGTGGTGGCGGATGCTGGGG 188
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1004 AGGTCTAATAGAGGTTTGGAGTCAAAATGCTTAAACTGCT---GGCAACATATAATAA 1059
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 187 AGAGTCTAATAGAGGTTTGGAGTCAAAATGCTTAAACTGCT---GGCAACATATAATAA 128
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1060 ATGCATGCTATTCAATGAATTTCTGCTATGAGGCAATCTGGCCCTGAGGAGCTCTC 1119
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 127 ATGCATGCTATTCAATGAATTTCTGCTATGAGGCAATCTGGCCCTGAGGAGCTCTC 77
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1120 CAGAAATTTCTAGTGAATTTCTGCTCTCTCTGCTTCAATGTTCTAATAACTTCTACATTAACA 1179
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 76 CAGAAATTTCTAGTGAATTTCTGCTCTCTCTGCTTCAATGTTCTAATAACTTCTACATTAACA 17
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1180 AAAAAAAAAAAAAA 1195
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 16 AAAAAAAAAAAAAA 1
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 47
BF523627
LOCUS
DEFINITION
  504 bp mRNA linear EST 11-DEC-2000
  UI-R-C1-1c-e-10-0-UI.r1 UI-R-C1 Rattus norvegicus cdna clone
  UI-R-C1-1c-e-10-0-UI 5', mRNA sequence.
ACCESSION
  BF523627
VERSION
  BF523627.1 GI:11631594
KEYWORDS
  EST.
SOURCE
  Rattus norvegicus (Norway rat)
  Rattus norvegicus
  Rattus norvegicus
REFERENCE
  1 (bases 1 to 504)
  Bonaldo,M.F., Lemmon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
JOURNAL
  97044477
MEDLINE
  8889548
PUBMED
  8889548
COMMENT
  Contact: Soares, MB
  Coordinated Laboratory for Computational Genomics
  University of Iowa
  375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
  Tel: 319 335 8250
  Fax: 319 335 9565
  Email: bento-soares@uiowa.edu
  cdna Library Preparation: M.B. Soares Lab Clone distribution:
  clones will be available through Research Genetics (www.resgen.com)
  This clone is also available through the I.M.A.G.E. Consortium at
  LNL (info@image.lnl.gov). IMAGE ID= 1792597
  Seq primer: M13 Forward.
  Location/Qualifiers
  1..504
  /organism="Rattus norvegicus"

```

```

/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C1-1c-e-10-0-UI"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-C1"
/note="Vector: p7713D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-C1
library is a subtracted library derived from the UI-R-C0
library, which is a subtracted library derived from the
UI-R-A1 and UI-R-E1 libraries. The UI-R-A1 library
consisted of a mixture of individually tagged normalized
libraries constructed from rat placenta, adult lung,
brain, liver, kidney, heart, spleen, ovary, and muscle.
The UI-R-E1 library consisted of a mixture of
individually tagged normalized libraries constructed from
8, 12 and 18-day embryo. The tag is a string of 3-5
nucleotides present between the Not I site and the
oligo-dT track which allows identification of the library
of origin of a clone within the mixture. The subtracted
library (UI-R-C1) was constructed as follows: PCR
amplified cDNA inserts from UI-R-C0 clones from which 3'
ESTs had been derived was used as a driver in a
hybridization with the UI-R-C0 library in the form of
single-stranded circles. The remaining single-stranded
circles (subtracted library) was purified by
hydroxyapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (Life Technologies) to generate the UI-R-C1
library. This procedure has been previously described
(Bonaldo, Lemmon and Soares, Genome Research 6: 791-806,
1996)."

```

ORIGIN

```

Query Match      30.6%; Score 366.2; DB 10; Length 504;
Best Local Similarity 86.5%; Pred. No. 3.5e-60;
Matches 444; Conservative 0; Mismatches 53; Indels 16; Gaps 3;

QY 668 GAAGATCTTCTCTTCCTGCCAACGAAAAAAGGGATTGAACAAAATGAACAGTGGTG 727
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4 GAGGATCTTCTCTCTTCCTACCAGCGAAAAAAGGGATTGACCAATGAGCAATGGGTG 63
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 728 GTCCCTCAAGTGAAGTAGAAGACCCGTCACGCCAGAACGAAGTGAAGAACTT 787
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 64 GTCCCAAGTGAAGTGAAGACCCGTCACGCCAGAACGAAGTGAAGAACTT 121
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 788 CCAATAAATGACTATCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 847
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 122 CCTGTTAATGACTATCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 181
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 848 TATGTTGTTATTTACTGCGTCCGAGCAACCGTATTTGCGCGCCGCTCTGTGAACCTTTA 907
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 182 TACTGTTGTTATTTACTGCGTCCGAGCAACCGTCTACTGCGCGAGGTCTGTGAACCTTTA 241
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 908 CTAGGCTTACTACCATATCCATCTGCTACCAAGGAGGACGAGTCTCTCTGTCATC 967
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 242 CTAGGCTTACTACCATATCCATCTGCTACCAAGGAGGTCGAGTCTCTCTGTCATC 301
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 968 ATGCTTTGTAATCTGGTGGTGGCGCCGCTGCTGGGAGGCTCTAATAGGAGTTTGAGCT 1027
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 302 ATGCTTTGTAATCTGGTGGTGGCGCCGCTGCTGGGAGGTCCTAATAGGAGTTTGAGCT 361
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1028 CAAATGCTTAACTGCT----GGCAACATATAATAATGCTATCTTCAATCAATTTCT 1083
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 362 CAAATGCTTAACTGCTTTTGTAGCAACATATAATTAATGCTGCTACTCTCAATTTCT 421
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1084 GCCTATGAGCATCTGGCCCTGTTAGGAGGCTCTCCAGAAATTTCTGTAGGTAATTCCT 1143
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 422 GCAT-----TTGCTCTCAAGTAGCCTATCTCTCCAGAAATTTATTGTAGGATATTCCT 471
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1144 CTCTTCATGTTCTAATAAACTTCTACATTAATCA 1176
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db      472 CTCTTCGTGTTCTTAATAAACGCTACATCATCA 504

RESULT 48
BI535434
LOCUS      404 bp mRNA linear EST 30-AUG-2001
DEFINITION
ACCESSION BI535434
VERSION    BI535434.1 GI:15376542
KEYWORDS   EST.
SOURCE     Bos taurus (cow)
ORGANISM   Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.
1 (bases 1 to 404)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.B., White,J., Cho,J., Fahrenkrug,S.C.,
Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
Chitko-Mckown,C.G., Perlea,G., Holt,I., Karamycheva,S., Liang,F.,
Quackenbush,J., and Keele,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
11282978
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCAGCAG
Plate: 127 row: B column: 3
Seq primer: ATTAGGTGACATATAG.
Location/Qualifiers
1. .404
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="WARC 4BOV"
/notes="Vector: pQMV SPORTE6; Site 1: Not1; Site 2: SalI;
Library made from pooled tissue from day 20 and day 40
embryos."

ORIGIN
Query Match 30.1%; Score 360.8; DB 12; Length 404;
Best Local Similarity 93.3%; Pred. No. 4.1e-59;
Matches 377; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY      128 AAATATGTAATCATTAAAGTTTGGAGTGGTATTCCTCGCCCTTAAGTCTA 187
Db      1 AAGATATGTAATTAAGTTTGGAGTGGTATTCCTCGCCCTTAAGTCTA 60

QY      188 ATTGTCCTGTTTGGGGGAGCAAGCAGCTTCTGGCCGAGGTACCCAAAAAGCCCTATGAC 247
Db      61 ATGTGTCCTGTTTGGGGGAGTAAAGCCTTCTGCTCTGAGACACCCAAAAACATATGAC 120

QY      248 ATGGAGCAGCATTCTTACAGCAATGAGAGAGAGAGATTACATGGAATTCATCTT 307
Db      121 ATGGAGCAGCATTCTTACAGCAATGAGAGAGAGAGATTACATGGAATTCATCTT 180

QY      308 GTGACAGCAATCGAATATTCAGACCGAATGCGACTGATGAACATTTGGAAGTGCAC 367
Db      181 ATTACAAAACCTGAATATTCAGAGTGAATGGCAGCTGATGAACATTTGGAAGTACAT 240

```

```

QY      368 GACTTTAAAAACGATACACATGGGATCTACTCTGTTGGTCTTCAAAAATGTTTATCAA 427
Db      241 GACTTTAAAAATGGATACATACACTGGCAATTTACTTTTGGAGTCTTCAAAAATGTTTATCAA 300

QY      428 ACTCAGATTAAAGTGATTCCCTGAATTTTCTGAACCAAGAGAGAAATAGATGAGATGAA 487
Db      301 ACTCAGATTAAAGTGATTCCCTGAATTTTCTGAACCAAGAGAGAAATAGATGAGATGAA 360

QY      488 GAAATATACCACAACCTTTCTTTGAAACAGTCAGTGATTTGGGTCCC 531
Db      361 GAAATATACCACAACCTTTCTTTGAAACAGTCAGTGATTTGGGTCCC 404

AI502787      492 bp mRNA linear EST 11-MAR-1999
UI-R-CI-1c-e-10-0-UI.s1 UI-R-CI Rattus norvegicus cDNA clone
UI-R-CI-1c-e-10-0-UI 3', mRNA sequence.
AI502787      GI:4400638
VERSION      AI502787.1
KEYWORDS     EST.
SOURCE       Rattus norvegicus (Norway rat)
ORGANISM     Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 492)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized 18 day embryo library cDNA Library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com)
Seq primer: M13 Forward.
Location/Qualifiers
1. .492
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CI-1c-e-10-0-UI"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-CI"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CI
library is a subtracted library derived from the UI-R-CI
library, which is a subtracted library derived from the
UI-R-AI and UI-R-EI libraries. The UI-R-AI library
consisted of a mixture of individually tagged normalized
libraries constructed from rat placenta, adult lung,
brain, liver, kidney, heart, spleen, ovary, and muscle.
The UI-R-EI library consisted of a mixture of
individually tagged normalized libraries constructed from
8, 12 and 18-day embryo. The tag is a string of 3-5
nucleotides present between the Not I site and the
oligo-dT track which allows identification of the library
of origin of a clone within the mixture. The subtracted

```


Library (UI-R-C1) was constructed as follows: PCR amplified cDNA inserts from UI-R-CO clones from which 3' ESTs had been derived were used as a driver in a hybridization with the UI-R-CO library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C1 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)." source

ORIGIN

```

Query Match      30.0%; Score 358.8; DB 9; Length 492;
Best Local Similarity 86.5%; Pred. No. 9.2e-59;
Matches 436; Conservative 0; Mismatches 52; Indels 16; Gaps 3;

QY 693 AAAAAAAGGATTGAACAAATGACAGTGGTGGTCCCTCAAGTGAAGTAGAGAAGA 752
Db      |||||
QY 492 AAAAAAAGGATTGACCAATGAGCAATGGTGGTCCCAACAGTGAAGTGGAGAAGA 433
Db      |||||
QY 753 CCGGTCACGCCAGACACAGCAAGTGAAGGAGACTTCCAAATGACTATCTGAAATG 812
Db      |||||
QY 432 CCGCGCCGACACAGACAGCAA--GAGGAAGACCTTCCTGTTAATGACTATCTGAAATG 375
Db      |||||
QY 813 GAATAGAAATTTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTTTACTGCCGTCGAG 872
Db      |||||
QY 374 GAATCGAATTTGATCCCATGCTGGATGAGAGAGGTTACTGTTGTTTACTGCCGTCGAG 315
Db      |||||
QY 873 GCAACCGCTATTGCCCGCGCTGTGTAACCTTTACTAGGCTACTACCATATCCATCTACT 932
Db      |||||
QY 314 GCAACCGCTACTGCCCGAGGCTGTGTAACCTTTACTAGGCTACTACCATATCCATCTACT 255
Db      |||||
QY 933 GCTACCAAGGAGGACGAGTCACTGTCGTGTCATCATGCTGCTGTAAGTGGTGGGCC 992
Db      |||||
QY 254 GCTACCAAGGAGTGGAGTCACTGTCGTGTCATCATGCTGCTGTAAGTGGTGGGCC 195
Db      |||||
QY 993 GCATGCTGGGAGGCTCTAATAGGAGGTTTGGAGTCAAACTGCTAACTGCT----GGCA 1048
Db      |||||
QY 194 GCATGCTGGGAGAGTCTAATAGGAGGTTTGGAGTCAAACTGCTAACTGCTTTGTAGCCA 135
Db      |||||
QY 1049 ACATATAATTAATGCTATCTAATGAATTTCTGCTTCTGAGGCACTGCGCCCTGGT 1108
Db      |||||
QY 134 ACATATAATTAATGCTATCTAATGAATTTCTGCTTCTGAGGCACTGCGCCCTGGT 85
Db      |||||
QY 1109 AGCCAGCTCTCCAGAAATTTACTGTTAGTAAATTTCTCTCTTCTCATGTTCTAATAAATCTCTA 1168
Db      |||||
QY 84 AGCTATCTCTCCAGAAATTTTGTAGGAAATTTCTCTCTTCTGTTTCTAATAAATCTCTA 25
Db      |||||
QY 1169 CATATACCAAAAAAATAAAAAA 1192
Db      |||||
QY 24 CATATCAAAAAAATAAAAAA 1
Db      |||||

```

```

RESULT 50
BF679243
LOCUS
DEFINITION 602153429F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4294374 5',
mRNA sequence.
ACCESSION BF679243
VERSION BF679243.1 GI:11953138
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 882)
NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-femail.nih.gov

```

Tissue Procurement: CLONETECH Laboratories, Inc.
 cDNA Library Preparation: CLONETECH Laboratories, Inc.
 cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCML143 row: j column: 07
 High quality sequence stop: 436.
 Location/Qualifiers
 1..882
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4294374"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH MGC 83"
 /note="Organ: prostate; Vector: pDNR-LiB (Clontech);
 Site 1: Sfil (ggcgctcgcc); Site 2: Sfil
 (ggcattatggcc); 5' and 3' adaptors were used in cloning
 as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'
 and 3' adaptor sequence:
 5'-ATTCTAGAGCGCGCGCCGACATG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.4
 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA)."

FEATURES

source

ORIGIN

```

Query Match      29.9%; Score 358.4; DB 10; Length 882;
Best Local Similarity 87.7%; Pred. No. 9.1e-59;
Matches 556; Conservative 0; Mismatches 51; Indels 27; Gaps 14;

QY 1 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGAAAGTAGTACTGTGCTCAGAGACCAT 60
Db      |||||
QY 27 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGAAAGTAGTACTGTGCTCAGAGACCAT 86
Db      |||||
QY 61 GGCAAGAATCTTCAGAGAATTTGTGAAGTGTGCATTTCTAAATGAGAAGCTTTTAA 120
Db      |||||
QY 87 GGC--AAGAATCTTCAGAGAATTTGTGAAGTGTGCATTTCTAAATGAGAAGCTTTTAA 145
Db      |||||
QY 121 ATCCAAAGAAATATGTAATCATTAAAGATTTGTGAAGTGTGCATTTCTAAATGAGAAGCTTTTAA 120
Db      |||||
QY 146 ATCCAAAGAAATATGTAATCATTAAAGATTTGTGAAGTGTGCATTTCTAAATGAGAAGCTTTTAA 205
Db      |||||
QY 181 AACTCTAATTTGCTCTGTTTGGGGGAGCAAGCACTTCTGGCCGAGGTACCCAAAAAAGC 240
Db      |||||
QY 206 AACTCTAATTTGCTCTG--TTTGGGGGAGCAAGCACTTCTGGCCGAGGTACCCAAAAAAGC 264
Db      |||||
QY 241 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAAGAAGATTTACAT--GGAAA 299
Db      |||||
QY 265 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAAGAAGATTTACATGGGAAA 324
Db      |||||
QY 300 TTGATCTGTGACCAAGCAAGTGAATATTTCAGAGCGGAATGGCACTGTGTAACATTTGG 359
Db      |||||
QY 325 TTGATCTGTGACCAAGCAAGTGAATATTTCAGAGCGGAATGGCACTGTGTAACATTTGG 384
Db      |||||
QY 360 --AAGTGCAGCACTTTTAAA--AAGCGATACACTGGCATCTACT--TCGTGGGTCTTTCAAAA 415
Db      |||||
QY 385 AAAGTGCAGCACTTTTAAAACACCGATACCACTGGCATCTACTCTCGTGGGTCTTTCAAAA 444
Db      |||||
QY 416 ----TGTTTATCAAAACTCAGATT--AAAGTGAATTCCTGAATTT--TCTGAAACCAAGA 468
Db      |||||
QY 445 CATGCTCTTATCACACACCTCAGATTACAGATGATTTCTGAAATTTAACTGAACACAGACAG 504
Db      |||||
QY 469 GGAAATA-----GATGAGAATGAAGAAATACCAACACTTTCTTTGA--ACAGTCACTGA 521
Db      |||||
QY 505 AGGAACACTAGATGAAGACTGACAGAAACTTACCAAACTTTCTTTTGACACAGTCACTGA 564
Db      |||||
QY 522 TTTGGGTCCAGCAGAAAAAGCCCTATTGAAAAACCGAGATTTTCTTAAAAAATTCCAAATTC 581
Db      |||||
QY 565 TT--GGGTCCAGCAGAAAGGGCTAT--GAAAACGAGATTTTCTTAAATTCAGAAATTC 620
Db      |||||

```

Tissue Procurement: CLONETECH Laboratories, Inc.

cDNA Library Preparation: CLONETECH Laboratories, Inc.

cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCML143 row: j column: 07

High quality sequence stop: 436.

Location/Qualifiers

1..882

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4294374"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NIH MGC 83"

/note="Organ: prostate; Vector: pDNR-LiB (Clontech);

Site 1: Sfil (ggcgctcgcc); Site 2: Sfil

(ggcattatggcc); 5' and 3' adaptors were used in cloning

as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'

and 3' adaptor sequence:

5'-ATTCTAGAGCGCGCGCCGACATG-dt(30)BN-3' (where B = A,

C, or G and N = A, C, G, or T). Average insert size 1.4

kb (range 0.5-4.0 kb). 14/15 colonies contained inserts

by PCR. This library was enriched for full-length clones

and was constructed by Clontech Laboratories (Palo Alto,

CA)."

Qy 582 TGGAG--ATTGTGATACGTGACCATGTATTGG 613
||| |
Db 621 TGGCGAATGGTGACAAACCGTGACCCGTATTGG 654
||| |

Search completed: September 3, 2004, 09:20:16
Job time : 2377 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 3, 2004, 04:35:02 ; Search time 3269 Seconds
(without alignments)
15870.790 Million cell updates/sec

Title: US-10-063-730-115

Perfect score: 1197

Sequence: 1 cagcagtggtctctcagtcctc.....caaaaaaaaaaaaaaaaa 1197

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1197	100.0	1197	6	AX092384	Sequence
2	1197	100.0	1197	6	AX464390	Sequence
3	1197	100.0	1197	6	AX697253	Sequence
4	1197	100.0	1197	9	AY358706	Homo sapi
5	1196	99.9	1228	6	BD269275	33 human
6	1194	99.7	1206	6	AX203127	Sequence
7	1187.2	99.2	1200	6	BD094021	A novel p
8	1178	98.4	1178	6	BD228713	Mammalian
9	1174	98.1	1380	6	AR338896	Sequence
10	1173.4	98.0	1214	9	AF234259	Homo sapi
11	1173	98.0	1309	9	AF191770	Homo sapi
12	1170	97.7	1184	9	AF291656	Homo sapi
13	1138.4	95.1	1140	9	AB055421	Homo sapi
14	1116	93.2	1428	6	AR338895	Sequence
15	1043.6	87.2	1175	4	AB059407	Equus cab
16	994.6	83.1	1338	10	BC049944	Mus muscu
17	994.6	83.1	1338	10	AF291655	Mus muscu
18	982.4	82.1	1202	10	AF219993	Mus muscu
19	979.2	81.8	1180	6	BD094022	A novel p
20	975	81.5	1197	6	BD094023	A novel p
21	964	80.5	1336	10	AF191769	Rattus no
22	961.8	80.4	1153	10	BC006919	Mus muscu
23	949.4	79.3	951	6	BD077398	Novel pro
24	949.4	79.3	951	6	BD096254	Novel pro
25	949.4	79.3	975	6	BD094032	Novel pro
26	943.6	78.8	1128	10	AB055422	Mus muscu
27	932.6	77.9	1214	10	AF191768	Mus muscu
28	926.2	77.4	1136	10	AB055423	Rattus no
29	823	68.8	951	6	AX203129	Sequence
30	823	68.8	975	6	BD094033	A novel p
31	805.4	67.3	807	6	BD077399	Novel pro
32	805.4	67.3	807	6	BD096255	Novel pro
33	378.8	31.6	97189	9	HS47937	Human DNA
34	377.2	31.5	20850	6	AX203136	Sequence
35	352.4	29.4	432	6	BD094036	A novel p
36	309	25.8	309	6	BD077400	Novel pro
37	309	25.8	309	6	BD096256	Novel pro
38	302	25.2	221461	10	AL691421	Mouse DNA
39	259	21.6	286	6	AX884585	Sequence
40	259	21.6	286	6	BD024195	Sequence
41	259	21.6	289	6	BD204197	5'EST and
42	247	20.6	248	6	BD228719	Mammalian
43	184.4	15.4	240491	2	AC107412	Rattus no
44	134	11.2	1347	10	BC045152	Mus muscu
45	130.8	10.9	1449	10	MM043509	Mus musculu

ALIGNMENTS

RESULT 1
AX092384
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AX092384
Sequence 115 from Patent WO0116318.
AX092384
AX092384.1 GI:13444506
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Eaton,D.L., Filvaroff,E., Gerritsen,M.E., Goddard,A.,
Godowski,P.J., Grimaldi,C.J., Gurney,A.L., Watanabe,C.K. and
Wood,W.I.

1197 bp DNA linear PAT 21-MAR-2001

TITLE Secreted and transmembrane polypeptides and nucleic acids encoding the same

JOURNAL Patent: WO 0116318-A 115 08-MAR-2001; Genentech, Inc. (US)

FEATURES Location/Qualifiers source

1.1197 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1197; DB 6; Length 1197; Best Local Similarity 100.0%; Pred. No. 4.2e-286; Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGAGAGAGTACTGTGCTGAGAGACCAT 60

DB 1 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGAGAGAGTACTGTGCTGAGAGACCAT 60

QY 61 GGCAGAGAAATCCCTCCAGAGAAATGTGAAGACTCTCACATTTCTAAATGCAGAACTTTTAA 120

DB 61 GGCAGAGAAATCCCTCCAGAGAAATGTGAAGACTCTCACATTTCTAAATGCAGAACTTTTAA 120

QY 121 ATCCAGAGAAATATGTAATACATTAAGATTTGTGAGTCTGTTGGTATCCTGSCCCT 180

DB 121 ATCCAGAGAAATATGTAATACATTAAGATTTGTGAGTCTGTTGGTATCCTGSCCCT 180

QY 181 AACTCTAAATGTCTCTGTTGGGGAGCAGACTCTTGCCCGGAGTACCCAAAAGC 240

DB 181 AACTCTAAATGTCTCTGTTGGGGAGCAGACTCTTGCCCGGAGTACCCAAAAGC 240

QY 241 CTATGACATGGAGCACACTTTCTACAGCAATGAGAGAGAGAAATTTACATGGAAAT 300

DB 241 CTATGACATGGAGCACACTTTCTACAGCAATGAGAGAGAGAAATTTACATGGAAAT 300

QY 301 TGATCTCTGACAGAACTGAATATTTCAAGACGGGAAATGGCACTGATGAACATTGGA 360

DB 301 TGATCTCTGACAGAACTGAATATTTCAAGACGGGAAATGGCACTGATGAACATTGGA 360

QY 361 AGTGCGAGCTTTTAAACCGGATACCTGCTGCTCTCTGCTGCTCTTCAAAAATGTTT 420

DB 361 AGTGCGAGCTTTTAAACCGGATACCTGCTGCTCTCTGCTGCTCTTCAAAAATGTTT 420

QY 421 TATCAAACTCAGATTAAGTGAATTCCTGAATTTCTGAACAGAGAGAAATAGATGA 480

DB 421 TATCAAACTCAGATTAAGTGAATTCCTGAATTTCTGAACAGAGAGAAATAGATGA 480

QY 481 GAATGAAGAAATTTACCAAACTTTCTTGAACAGTCACTGATTTGGTCCAGCAGAAA 540

DB 481 GAATGAAGAAATTTACCAAACTTTCTTGAACAGTCACTGATTTGGTCCAGCAGAAA 540

QY 541 GCCTATTGAAAACCGAGATTTCTTCAAAAATTCAAAATTTCTGAGATTTGTGATAACGT 600

DB 541 GCCTATTGAAAACCGAGATTTCTTCAAAAATTCAAAATTTCTGAGATTTGTGATAACGT 600

QY 601 GACCATGATTTGGATCAATCCACTCTTAATATCAATTTCTGAGTTAACAAGCTTTGAGGA 660

DB 601 GACCATGATTTGGATCAATCCACTCTTAATATCAATTTCTGAGTTAACAAGCTTTGAGGA 660

QY 661 GGAGGGAGAGATTTTCACTTTCTGCGCAAGAAAAGGGATTCGAACAATGAACA 720

DB 661 GGAGGGAGAGATTTTCACTTTCTGCGCAAGAAAAGGGATTCGAACAATGAACA 720

QY 721 GTGGGTGGTCCCTCAAGTGAAGTAGAGAGACCCGTCAAGCCAGACAAGCAAGTGAAGA 780

DB 721 GTGGGTGGTCCCTCAAGTGAAGTAGAGAGACCCGTCAAGCCAGACAAGCAAGTGAAGA 780

QY 781 AGAATTTCCAAATAATGATCTATGTAATTTGGAATTTGATCCCATGCTGGATGA 840

DB 781 AGAATTTCCAAATAATGATCTATGTAATTTGGAATTTGATCCCATGCTGGATGA 840

QY 841 GAGAGGTATTTGTTATTTTACTGCGTCGAGCAACCGCTATTGCGCGCGCTCTGTGA 900

DB 841 GAGAGGTATTTGTTATTTACTCGCTCGAGCAACCGCTATTGCGCGCGCTCTGTGA 900

QY 901 ACCTTTACTAGGCTACTACCCATATCCATATCTGCTACCAAGAGAGAGTATCTGTGCG 960

DB 901 ACCTTTACTAGGCTACTACCCATATCCATATCTGCTACCAAGAGAGAGTATCTGTGCG 960

QY 961 TGTCTATCATGCTTGTAACTGGTGGTGGCCCGCATGCTGGGGAGGGTCTAAATAGGAGGT 1020

DB 961 TGTCTATCATGCTTGTAACTGGTGGTGGCCCGCATGCTGGGGAGGGTCTAAATAGGAGGT 1020

QY 1021 TTGAGCTCAAAATGCTTTAAATCTCTGCAACATATATAATATGATCTATTTCAATGAAT 1080

DB 1021 TTGAGCTCAAAATGCTTTAAATCTCTGCAACATATATAATATGATCTATTTCAATGAAT 1080

QY 1081 TCTGCTATGAGGCACTCTGCCCCCTGGTAGCCAGCTCTCCAGAAATTTACTTTGAGGTAAT 1140

DB 1081 TCTGCTATGAGGCACTCTGCCCCCTGGTAGCCAGCTCTCCAGAAATTTACTTTGAGGTAAT 1140

QY 1141 CCTCTCTTCAATGTTCTAAATAAACTTTACATATATCACCAAAAAA 1197

DB 1141 CCTCTCTTCAATGTTCTAAATAAACTTTACATATATCACCAAAAAA 1197

RESULT 2

AX464390 1197 bp DNA linear PAT 16-JUL-2002

LOCUS Sequence 523 from Patent WO0140466.

DEFINITION AX464390

ACCESSION AX464390.1 GI:21899211

VERSION

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Baker, K.P., Beresini, M., DeForge, L., Desnoyers, L., Filvaroff, E., Gao, W.Q., Gerritsen, M.E., Goddard, A., Godowski, P.J., Gurney, A.L., Sherwood, S., Smith, V., Stewart, T.A., Tomas, D., Watanabe, C.K., Wood, W.L. and Zhang, Z.

TITLE Secreted and transmembrane polypeptides and nucleic acids encoding same

JOURNAL Patent: WO 0140466-A 523 07-JUN-2001; Genentech Inc. (US)

FEATURES Location/Qualifiers source

1.1197 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1197; DB 6; Length 1197; Best Local Similarity 100.0%; Pred. No. 4.2e-286; Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGAGAGAGTACTGTGCTGAGAGACCAT 60

DB 1 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGAGAGAGTACTGTGCTGAGAGACCAT 60

QY 61 GGCAGAGAAATCCCTCCAGAGAAATGTGAAGACTCTCACATTTCTAAATGCAGAACTTTTAA 120

DB 61 GGCAGAGAAATCCCTCCAGAGAAATGTGAAGACTCTCACATTTCTAAATGCAGAACTTTTAA 120

QY 121 ATCCAGAGAAATATGTAATACATTAAGATTTGTGAGTCTGTTGGTATCCTGSCCCT 180

DB 121 ATCCAGAGAAATATGTAATACATTAAGATTTGTGAGTCTGTTGGTATCCTGSCCCT 180

QY 181 AACTCTAAATGTCTCTGTTGGGGAGCAGACTCTTGCCCGGAGTACCCAAAAGC 240

DB 181 AACTCTAAATGTCTCTGTTGGGGAGCAGACTCTTGCCCGGAGTACCCAAAAGC 240

QY 241 CTATGACATGGAGCACACTTTCTACAGCAATGAGAGAGAGAAATTTACATGGAAAT 300

DB 241 CTATGACATGGAGCACACTTTCTACAGCAATGAGAGAGAGAAATTTACATGGAAAT 300

QY 301 TGATCCTGTGACGAGAACTGAAATATTGAGAGCGGAAATGGCACTGATGAAACATTGGA 360
Db 301 TGATCCTGTGACGAGAACTGAAATATTGAGAGCGGAAATGGCACTGATGAAACATTGGA 360
QY 361 AGTGCAGCACTTTAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTT 420
Db 361 AGTGCAGCACTTTAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTT 420
QY 421 TATCAAACTCAGATTAAAGTGATTCCTCAATTTCTGAACAGTCAAGTGGTCCCGAGCAAAA 480
Db 421 TATCAAACTCAGATTAAAGTGATTCCTCAATTTCTGAACAGTCAAGTGGTCCCGAGCAAAA 480
QY 481 GAATGAAGAAATTAACCAAACTTTCTTTGAAACAGTCAAGTGGTCCCGAGCAAAA 540
Db 481 GAATGAAGAAATTAACCAAACTTTCTTTGAAACAGTCAAGTGGTCCCGAGCAAAA 540
QY 541 GCCTATTGAAACCGAGATTTCTTAAAATTTCCAAAATTTGGGATTTGATGAAAGT 600
Db 541 GCCTATTGAAACCGAGATTTCTTAAAATTTCCAAAATTTGGGATTTGATGAAAGT 600
QY 601 GACCATGATTGGATCAATCCCACTTAATATCAGTTTCTGAGTTACAGACTTTGAGGA 660
Db 601 GACCATGATTGGATCAATCCCACTTAATATCAGTTTCTGAGTTACAGACTTTGAGGA 660
QY 661 GGAGGAGAAGATCTTCACTTTCTGCGCAACGAAAAAAGGATTTGAAACAAAATGAACA 720
Db 661 GGAGGAGAAGATCTTCACTTTCTGCGCAACGAAAAAAGGATTTGAAACAAAATGAACA 720
QY 721 GTGGTGTGCTCAAGTGAAGTGAAGAACCCGTCACGCGACAGCAAGTGAAGGA 780
Db 721 GTGGTGTGCTCAAGTGAAGTGAAGAACCCGTCACGCGACAGCAAGTGAAGGA 780
QY 781 AGAACTTCAATAAATGACTATCTGAAATGGAATGGAATTTGATTCCTGATGGAATGA 840
Db 781 AGAACTTCAATAAATGACTATCTGAAATGGAATGGAATTTGATTCCTGATGGAATGA 840
QY 841 GAGAGGTTATTGTTGATTTACTGCGGTGCGAGCAACCGCTATTGCGCGCGCTCTGTGA 900
Db 841 GAGAGGTTATTGTTGATTTACTGCGGTGCGAGCAACCGCTATTGCGCGCGCTCTGTGA 900
QY 901 ACCTTTACTAGGCTACTACCAATATCCATATCTGTTACCAAGGAGGACGAGTCACTGTCG 960
Db 901 ACCTTTACTAGGCTACTACCAATATCCATATCTGTTACCAAGGAGGACGAGTCACTGTCG 960
QY 961 TGTCAATCATGCTTGTAACTGGTGGTGGCGCGATGCTGGGAGGGTCTAATAGGAGT 1020
Db 961 TGTCAATCATGCTTGTAACTGGTGGTGGCGCGATGCTGGGAGGGTCTAATAGGAGT 1020
QY 1021 TTGAGCTCAATGCTTAACTGCTGGCAACATATAATAAATGCAATGCTATTCAATGAAT 1080
Db 1021 TTGAGCTCAATGCTTAACTGCTGGCAACATATAATAAATGCAATGCTATTCAATGAAT 1080
QY 1081 TCTGCTTATGAGGCTATGCGCCCTGTAGCAGGCTCTCCAGAAATTTCTGAGTAAT 1140
Db 1081 TCTGCTTATGAGGCTATGCGCCCTGTAGCAGGCTCTCCAGAAATTTCTGAGTAAT 1140
QY 1141 CCTCTCTTCAATGCTTAACTGCTGGCAACATATAATAAATGCAATGCTATTCAATGAAT 1197
Db 1141 CCTCTCTTCAATGCTTAACTGCTGGCAACATATAATAAATGCAATGCTATTCAATGAAT 1197

RESULT 3
LOCUS AX697253
DEFINITION Sequence 321 from Patent W00078961.
ACCESSION AX697253
VERSION AX697253.1 GI:29498415
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN

1 Ferrara, N., Stewart, T.A., Williams, P.M., Baker, K.P., Desnoyers, L.,
Eaton, D.L., Gao, W.Q., Pan, J., Botstein, D., Fong, S., Goddard, A.,
Godowski, P.J., Gurney, A.L., Smith, V., Tumas, D., Wood, W.I.,
Grimaldi, C.J., Hillan, K.J., Paoni, N.F., Roy, M.A. and Watanabe, C.K.
Secreted and transmembrane polypeptides and nucleic acids encoding
the same
Patent: WO 0078961-A 321 28-DEC-2000;
Genentech Inc. (US)
Location/Qualifiers
1..1197
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
Query Match 100.0%; Score 1197; DB 6; Length 1197;
Best Local Similarity 100.0%; Pred. No. 4.2e-286;
Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGAAAGTACTGTGCTGAGAGCAAT 60
Db 1 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGAAAGTACTGTGCTGAGAGCAAT 60
QY 61 GCGAAAGAAATCCTCCAGAGAATCTGGAAGCTGTGCACATTTCTAATGCAAGCTTTTAA 120
Db 61 GCGAAAGAAATCCTCCAGAGAATCTGGAAGCTGTGCACATTTCTAATGCAAGCTTTTAA 120
QY 121 ATCCAAGAAAAATATGTAATCACTTAAGATTTGTGGACTGTGTTGCTGCTGGCCCT 180
Db 121 ATCCAAGAAAAATATGTAATCACTTAAGATTTGTGGACTGTGTTGCTGCTGGCCCT 180
QY 181 AACTCTAATTTGCTGTTTTGGGGAGCAAGCACTTTCTGGCGGAGGTACCCAAAAAGC 240
Db 181 AACTCTAATTTGCTGTTTTGGGGAGCAAGCACTTTCTGGCGGAGGTACCCAAAAAGC 240
QY 241 CTATGACATGAGACACACTTTCTACAGCAATGGAGAGAAAGAGATTACATGGAAT 300
Db 241 CTATGACATGAGACACACTTTCTACAGCAATGGAGAGAAAGAGATTACATGGAAT 300
QY 301 TGAATCCTGTGACAGAACTGAAATATTGAGAGCGGAAATGGCACTGATGAAACATTGGA 360
Db 301 TGAATCCTGTGACAGAACTGAAATATTGAGAGCGGAAATGGCACTGATGAAACATTGGA 360
QY 361 AGTGCAGCACTTTAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTT 420
Db 361 AGTGCAGCACTTTAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTT 420
QY 421 TATCAAACTCAGATTAAAGTGATTCCTCAATTTCTGAACAGTCAAGTGGTCCCGAGCAAAA 480
Db 421 TATCAAACTCAGATTAAAGTGATTCCTCAATTTCTGAACAGTCAAGTGGTCCCGAGCAAAA 480
QY 481 GAATGAAGAAATTAACCAAACTTTCTTTGAAACAGTCAAGTGGTCCCGAGCAAAA 540
Db 481 GAATGAAGAAATTAACCAAACTTTCTTTGAAACAGTCAAGTGGTCCCGAGCAAAA 540
QY 541 GCCTATTGAAACCGAGATTTCTTAAAATTTCCAAAATTTGGGATTTGATGAAAGT 600
Db 541 GCCTATTGAAACCGAGATTTCTTAAAATTTCCAAAATTTGGGATTTGATGAAAGT 600
QY 601 GACCATGATTGGATCAATCCCACTTAATATCAGTTTCTGAGTTACAGACTTTGAGGA 660
Db 601 GACCATGATTGGATCAATCCCACTTAATATCAGTTTCTGAGTTACAGACTTTGAGGA 660
QY 661 GGAGGAGAAGATCTTCACTTTCTGCGCAACGAAAAAAGGATTTGAAACAAAATGAACA 720
Db 661 GGAGGAGAAGATCTTCACTTTCTGCGCAACGAAAAAAGGATTTGAAACAAAATGAACA 720
QY 721 GTGGTGTGCTCAAGTGAAGTGAAGAACCCGTCACGCGACAGCAAGTGAAGGA 780
Db 721 GTGGTGTGCTCAAGTGAAGTGAAGAACCCGTCACGCGACAGCAAGTGAAGGA 780
QY 781 AGAACTTCAATAAATGACTATCTGAAATGGAATGGAATTTGATTCCTGATGGAATGA 840

Db 781 AGAATCTCCAAATAGTACTACTGAAATGGAATAGAAATTTGATCCCATGCTGATGA 840
QY 841 GAGAGGTTATTTGTTATTTAATGCGGTGAGAGCAACCCCTATTTGCCCGCGGTCTGTGA 900
Db 841 GAGAGGTTATTTGTTATTTAATGCGGTGAGAGCAACCCCTATTTGCCCGCGGTCTGTGA 900
QY 901 ACCTTTACTAGGCTACTACCCATATCCATCTACTGCTACCAAGAGGACGAGTCTATCTGTCG 960
Db 901 ACCTTTACTAGGCTACTACCCATATCCATCTACTGCTACCAAGAGGACGAGTCTATCTGTCG 960
QY 961 TGTCTATCATGCTTGTAACTGTTGGTGGTGGCCCGCATGCTGTTGGGAGGTTCTAATPAGAGGT 1020
Db 961 TGTCTATCATGCTTGTAACTGTTGGTGGTGGCCCGCATGCTGTTGGGAGGTTCTAATPAGAGGT 1020
QY 1021 TTGAGGTCGAATGCTTAACTGCTGCAACATATAATAATGATGCTATTCGAATCAAT 1080
Db 1021 TTGAGGTCGAATGCTTAACTGCTGCAACATATAATAATGATGCTATTCGAATCAAT 1080
QY 1081 TCTGCTATGAGGCTATGCGCCCTGTTAGCCAGCTCTCCAGAAATTTACTTTGTAGGTAAT 1140
Db 1081 TCTGCTATGAGGCTATGCGCCCTGTTAGCCAGCTCTCCAGAAATTTACTTTGTAGGTAAT 1140
QY 1141 CCTCTCTCATGTTCTTAATAAATTTCTACATTATACCAAAAAA 1197
Db 1141 CCTCTCTCATGTTCTTAATAAATTTCTACATTATACCAAAAAA 1197
RESULT 4
AY358706 1197 bp mRNA linear PRI 03-OCT-2003
LOCUS Homo sapiens clone DNA73727 AKNP771 (UNQ771) mRNA, complete cds.
DEFINITION AY358706
ACCESSION AY358706.1 GI:37182533
VERSION FLI_CDNA.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM
REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1197)
Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J., Chen,J., Chow,B., Chui,C., Crowley,C., Currell,B., Deuel,B., Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.E., Heldens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S., Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C., Schoenfeld,J., Sehagiri,S., Simmons,L., Singh,J., Smith,V., Stinson,J., Vagts,A., Vandlen,R., Watanabe,C., Wieand,D., Woods,K., Xie,M.H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z., Goddard,A., Wood,W.I. and Godowski,P.
The Secreted Protein Discovery Initiative (SPDI), a Large-Scale Effort to Identify Novel Human Secreted and Transmembrane Proteins: A Bioinformatics Assessment
Genome Res. 13 (10), 2265-2270 (2003)
JOURNAL
PUBMED 12975309
REFERENCE 2 (bases 1 to 1197)
AUTHORS Clark,H.F.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-2003) Department of Bioinformatics, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA
FEATURES
source
1. .1197
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DNA73727"
1. .1197
/locus_tag="UNQ771"
59. .1012
/locus_tag="UNQ771"
/note="PRO1565"
/codon_start=1
/product="AKNP771"
/protein_id="AAQ89069.1"

/db_xref="GI:37182534"
/translation="MAKPPENCEDCHILNIAEAFKSKIKCKLTKGLVFLALTLI
VLFKNGHFWPEPKAYDMETHFYNGEKKIYMBIDFVTRTEIFRSNGDETLIV
HDFKNGYGTGIYFVGLQKCFIKTIKVIPEFSEPEEIDENEETITTFPFQSWIWPAA
KPIENRDLKNSKILEICDNVTMYNPTLISVSELQPEEEDLHPFPAKKGIEQ
NEQWTVQVKVEKTEHARQASEEELPINDYTENGIEFDPMLEDRGVYCCIIYCRNRNYC
RRVCEPLLGLTYPIFYCYGGGRVIVCRVIMPCNWWARMGLRV"
ORIGIN
Query Match 100.0%; Score 1197; DB 9; Length 1197;
Best Local Similarity 100.0%; Pred. No. 4.2e-286;
Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGCAGTGTCTCTCAGTCTCTCTAAAGCAAGAAAGAGTACTGTCTGAGAGACCAT 60
Db 1 CAGCAGTGTCTCTCAGTCTCTCTAAAGCAAGAAAGAGTACTGTCTGAGAGACCAT 60
QY 61 GGCAGAAATCTCCAGAGAAATGTGAAGACTGTCACTTTAAATGCGAAGCTTTTAA 120
Db 61 GGCAGAAATCTCTCCAGAGAAATGTGAAGACTGTCACTTTAAATGCGAAGCTTTTAA 120
QY 121 ATCCAGAAAAATATGTAAATCACTTAAGATTTGTGCACTGGTGTGGTATCTGGCCCT 180
Db 121 ATCCAGAAAAATATGTAAATCACTTAAGATTTGTGCACTGGTGTGGTATCTGGCCCT 180
QY 181 AACTCTAATGTCTGTTTGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAGC 240
Db 181 AACTCTAATGTCTGTTTGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAGC 240
QY 241 CTATGACATGGAGCACACTTTCTACAGCAATGAGAGAGAAAGATTTACATGGAAT 300
Db 241 CTATGACATGGAGCACACTTTCTACAGCAATGAGAGAGAAAGATTTACATGGAAT 300
QY 301 TGATCTCTGTGACCAAGCACTGAAATATTTCAGAACGGGAAATGGCACTGATGAATGGA 360
Db 301 TGATCTCTGTGACCAAGCACTGAAATATTTCAGAACGGGAAATGGCACTGATGAATGGA 360
QY 361 AGTGCGACGCTTTAAAAACGGATACACTGGCATCTACTTCTGGTGGTCTTCAAAAATGTTT 420
Db 361 AGTGCGACGCTTTAAAAACGGATACACTGGCATCTACTTCTGGTGGTCTTCAAAAATGTTT 420
QY 421 TATCAAACTCAGATTAAGTGATTTCTGAATTTTCTGAACCAAGAGAGAAATAGATGA 480
Db 421 TATCAAACTCAGATTAAGTGATTTCTGAATTTTCTGAACCAAGAGAGAAATAGATGA 480
QY 481 GAATGAAGAAATATACCAAACTTTCTTTGAACAGTCACTGATTTGGTCCAGCAGAAAA 540
Db 481 GAATGAAGAAATATACCAAACTTTCTTTGAACAGTCACTGATTTGGTCCAGCAGAAAA 540
QY 541 GCCTATTGAAAAACGAGATTTCTTAAAAATTTCCAAAATTTCTGGAGATTTGTGATAACGT 600
Db 541 GCCTATTGAAAAACGAGATTTCTTAAAAATTTCCAAAATTTCTGGAGATTTGTGATAACGT 600
QY 601 GACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAGGA 660
Db 601 GACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAGGA 660
QY 661 GGAGGGAGAGATCTTCACTTTCTGCCAACGAAAAAAGGGATGAAACAAAATGAGCA 720
Db 661 GGAGGGAGAGATCTTCACTTTCTGCCAACGAAAAAAGGGATGAAACAAAATGAGCA 720
QY 721 GTGGGTGGTCCCTCAAGTGAAGTAGAGAGACCGGTACGCCAGACAGCAAGTGAAGGA 780
Db 721 GTGGGTGGTCCCTCAAGTGAAGTAGAGAGACCGGTACGCCAGACAGCAAGTGAAGGA 780
QY 781 AGAATCTCCAAATAATGACTATCTGAAAAATGGAATAGAAATTTGATCCCATGCGATGA 840
Db 781 AGAATCTCCAAATAATGACTATCTGAAAAATGGAATAGAAATTTGATCCCATGCGATGA 840
QY 841 GAGAGGTTATTTGTTATTTACTGCCGTGAGAGCAACCGCTATTTCGCCCGCGGTCTGTGA 900
Db 841 GAGAGGTTATTTGTTATTTACTGCCGTGAGAGCAACCGCTATTTCGCCCGCGGTCTGTGA 900


```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1200)
Yamana,K., Takahashi,Y., Wada,H. and Kasahara,Y.
A novel polypeptide and its encoding gene
Patent: WO 0123557-A 1 05-APR-2001;
TEIJIN LTD,KEI YAMANA,YUKIMI TAKAHASHI,HITOSHI WADA, YOSHINORI
KASAHARA
OS Homo sapiens (human)
PN WO 0123557-A/1
PD 05-APR-2001
PF 29-SEP-2000 WO 2000JP006904
PR 29-SEP-1999 JP 99P 275947
PI KEI YAMANA,YUKIMI TAKAHASHI,HITOSHI WADA,YOSHINORI KASAHARA PC
C12N15/12,C12Q1/68,C12P21/08,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
10
CC
FH Key Location/Qualifiers
FT CDS (67)..(1020)..
source
Location/Qualifiers
1..1200
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.2%; Score 1187.2; DB 6; Length 1200;
Best Local Similarity 99.7%; Pred. No. 1.1e-283;
Matches 1189; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGGAAGAGTACTGTGTGCTGAGAGCAAT 60
DB 9 CAGCAGGTGTCTCTCAGTCTCTCAAGCAAGGAAGAGTACTGTGTGCTGAGAGCAAT 68
QY 61 GCGAAGAAATCCTCCAGAGAAATCTGAGACTGTGCATTTCAATTCAGAGCTTTTAA 120
DB 69 GCGAAGAAATCCTCCAGAGAAATCTGAGACTGTGCATTTCAATTCAGAGCTTTTAA 128
QY 121 ATCCAGAAATATGTAAATCACCTTAAGATTTGTGGACTGGTGTTCCTGCGCCCT 180
DB 129 ATCCAGAAATATGTAAATCACCTTAAGATTTGTGGACTGGTGTTCCTGCGCCCT 188
QY 181 AACTCTAATTTGCTCTGTTTGGGGGAGCAAGCACTTCTGGCCGAGGTACCCAAAAAAGC 240
DB 189 AACTCTAATTTGCTCTGTTTGGGGGAGCAAGCACTTCTGGCCGAGGTACCCAAAAAAGC 248
QY 241 CTATGACATGAGCAGACACTTTCTACAGCAATGGAGAGAGAGATTTACATGGAAAT 300
DB 249 CTATGACATGAGCAGACACTTTCTACAGCAATGGAGAGAGAGATTTACATGGAAAT 308
QY 301 TGATCTGTGACCAGAACTGAAATATTCAGAAAGCGGAAATGGCACTGATGAAACATTGGA 360
DB 309 TGATCTGTGACCAGAACTGAAATATTCAGAAAGCGGAAATGGCACTGATGAAACATTGGA 368
QY 361 AGTGACGACCTTTAAACCGGATACACTGGCACTTCTCGTGGGCTTCAAAAATGTTT 420
DB 369 AGTACGACCTTTAAACCGGATACACTGGCACTTCTCGTGGGCTTCAAAAATGTTT 428
QY 421 TATCAAACTCAGATTAAGATTCCTGAAATTTCTGACCCAGAGAGAAATAGATGA 480
DB 429 TATCAAACTCAGATTAAGATTCCTGAAATTTCTGACCCAGAGAGAAATAGATGA 488
QY 481 GAATGAAGAAATACCACAACTTCTTTGAACAGTCACTGATTTGGGTCCAGCAGAAAA 540
DB 489 GAATGAAGAAATACCACAACTTCTTTGAACAGTCACTGATTTGGGTCCAGCAGAAAA 548
QY 541 GCCTATTGAAACCGGAGATTTCTTAAAAATTCAAAAATTCAGAGATTTGTGATACGT 600
DB 549 GCCTATTGAAACCGGAGATTTCTTAAAAATTCAAAAATTCAGAGATTTGTGATACGT 608
QY 601 GACCATGATTGGATCAATCCCACTCTAAATATCAGTTTCTGAGTTACAGACTTTGAGGA 660
DB 609 GACCATGATTGGATCAATCCCACTCTAAATATCAGTTTCTGAGTTACAGACTTTGAGGA 668
```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
OS Homo sapiens (human)
PN JP 2002530078-A/1
PD 17-SEP-2002
PF 12-NOV-1999 JP 2000582562
PR 13-NOV-1998 US 09/191986
PI SI LOK,SCOTT R PRESNELL
PC C12N15/09,C07K14/51,C07K16/24,C12N15/00
CC Mammalian chondromodulin-like protein
FH Key Location/Qualifiers
FT CDS (58)..(1008)..
source
Location/Qualifiers
1..1178
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 98.4%; Score 1178; DB 6; Length 1178;
Best Local Similarity 100.0%; Pred. No. 2.2e-281;
Matches 1178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 661 GGAGGAGAGATCTTCACTTTCTCGCCAAAGAAAAAGGATTGAACAAAATGAACA 720
DB 669 GGAGGAGAGATCTTCACTTTCTCGCCAAAGAAAAAGGATTGAACAAAATGAACA 728
QY 721 GTGGGTGGTCCCTCAAGTGAAGTAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA 780
DB 729 GTGGGTGGTCCCTCAAGTGAAGTAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA 788
QY 781 AGAACTTCCAAATAATGACTATCTGAAATGGAATAGAAATTTGATCCATCTGGATGA 840
DB 789 AGAACTTCCAAATAATGACTATCTGAAATGGAATAGAAATTTGATCCATCTGGATGA 848
QY 841 GAGAGGTATTGTTGTTATTTACTGCGTCGAGGCAACCGCTATTGCGCCGCTGCTGGA 900
DB 849 GAGAGGTATTGTTGTTATTTACTGCGTCGAGGCAACCGCTATTGCGCCGCTGCTGGA 908
QY 901 ACCTTTACTAGGCTACTACCCATATCCATATCTGCTACCAAGGAGAGAGTCACTGTGCG 960
DB 909 ACCTTTACTAGGCTACTACCCATATCCATATCTGCTACCAAGGAGAGAGTCACTGTGCG 968
QY 961 TGTCACTATGCTTGTAACTGGTGGTGGCCCGCATGCTGGGAGGGTCTTAATAGGAGGT 1020
DB 969 TGTCACTATGCTTGTAACTGGTGGTGGCCCGCATGCTGGGAGGGTCTTAATAGGAGGT 1028
QY 1021 TTGAGCTCAAAATGCTTAACTGCTGGCAACATATAATAATGTCATGCTATTCAATGAAT 1080
DB 1029 TTGAGCTCAAAATGCTTAACTGCTGGCAACATATAATAATGTCATGCTATTCAATGAAT 1088
QY 1081 TCTGCTTATGAGGCATCTGCCCCCTGTAGCAGCTCTCCAGAAATTTACTTGTAGTAAT 1140
DB 1089 TCTGCTTATGAGGCATCTGCCCCCTGTAGCAGCTCTCCAGAAATTTACTTGTAGTAAT 1148
QY 1141 CCTCTCTCATGTTCTTAATAAATTTCTACATTATCACAAAAAATAAAAAA 1192
DB 1149 CCTCTCTCATGTTCTTAATAAATTTCTACATTATCACAAAAAATAAAAAA 1200
RESULT 8
BD228713 1178 bp DNA linear PAT 17-JUL-2003
LOCUS Mammalian chondromodulin-like protein.
DEFINITION BD228713
ACCESSION BD228713
VERSION BD228713.1 GI:33038483
KEYWORDS JP 2002530078-A/1
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1178)
AUTHORS Lok,S. and Presnell,S.R.
TITLE Mammalian chondromodulin-like protein
JOURNAL Patent: JP 2002530078-A 1 17-SEP-2002;
ZYMOGENETICS INC
COMMENT OS Homo sapiens (human)
PN JP 2002530078-A/1
PD 17-SEP-2002
PF 12-NOV-1999 JP 2000582562
PR 13-NOV-1998 US 09/191986
PI SI LOK,SCOTT R PRESNELL
PC C12N15/09,C07K14/51,C07K16/24,C12N15/00
CC Mammalian chondromodulin-like protein
FH Key Location/Qualifiers
FT CDS (58)..(1008)..
source
Location/Qualifiers
1..1178
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
```

QY 2 AGCAGTGGTCTCTCAGTCTCTCAAGCAAGAAAGAGTACTGTGTGCTGAGAGACCATG 61
DB |||||
QY 1 AGCAGTGGTCTCTCAGTCTCTCAAGCAAGAAAGAGTACTGTGTGCTGAGAGACCATG 60
DB |||||
QY 62 GCAAGAATCCTCCAGAGAAATGTGAAGACTGTGCATCTCTAAATGCAGAGCTTTTAAA 121
DB |||||
QY 61 GCAAGAATCCTCCAGAGAAATGTGAAGACTGTGCATCTCTAAATGCAGAGCTTTTAAA 120
DB |||||
QY 122 TCCAGAAATATGTAATCACTTAAGATTGTGGACTGTGTGGTATCTCTGGCCCTA 181
DB |||||
QY 121 TCCAGAAATATGTAATCACTTAAGATTGTGGACTGTGTGGTATCTCTGGCCCTA 180
DB |||||
QY 182 ACTCTAAATGTCTCTGTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCC 241
DB |||||
QY 181 ACTCTAAATGTCTCTGTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCC 240
DB |||||
QY 242 TATGACATGGAGACACTTTCTACAGCAATGGAGAGAAAGAAATTTACATGGAATTT 301
DB |||||
QY 241 TATGACATGGAGACACTTTCTACAGCAATGGAGAGAAAGAAATTTACATGGAATTT 300
DB |||||
QY 302 GATCCTGTGACAGAACTCAAAATATTCAAGCGGAAATGGCACTGATGAACATTTGGA 361
DB |||||
QY 301 GATCCTGTGACAGAACTCAAAATATTCAAGCGGAAATGGCACTGATGAACATTTGGA 360
DB |||||
QY 362 GTGACGACTTTAAAAACGGATACACTGGCATCTCTCTGGGTCTTCAAAAATGTTTT 421
DB |||||
QY 361 GTGACGACTTTAAAAACGGATACACTGGCATCTCTCTGGGTCTTCAAAAATGTTTT 420
DB |||||
QY 422 ATCAAAACTCAGATTAAGTGATCTCTGAATTTCTGAACCAAGAGAAATAGATGAG 481
DB |||||
QY 421 ATCAAAACTCAGATTAAGTGATCTCTGAATTTCTGAACCAAGAGAAATAGATGAG 480
DB |||||
QY 482 AATGAAGAATTAACCAACTTTCTTTGAACAGTCAAGTGAATTTGGGTCCAGCAGAAAAAG 541
DB |||||
QY 481 AATGAAGAATTAACCAACTTTCTTTGAACAGTCAAGTGAATTTGGGTCCAGCAGAAAAAG 540
DB |||||
QY 542 CTTATTGAAAACCGAGATTTCTTAAAAATTCAAAAATTCGAGATTTGTGATAACGTG 601
DB |||||
QY 541 CTTATTGAAAACCGAGATTTCTTAAAAATTCAAAAATTCGAGATTTGTGATAACGTG 600
DB |||||
QY 602 ACCATGTTTGGATCAATCCCACTCTAATCAGTTTCTGAGTTACAAGACTTTGAGGAG 661
DB |||||
QY 601 ACCATGTTTGGATCAATCCCACTCTAATCAGTTTCTGAGTTACAAGACTTTGAGGAG 660
DB |||||
QY 662 GAGGAGAGATCTTCACTTTCTGCGCAACGAAAAAAGGGATTGAACAAAAATGAACAG 721
DB |||||
QY 661 GAGGAGAGATCTTCACTTTCTGCGCAACGAAAAAAGGGATTGAACAAAAATGAACAG 720
DB |||||
QY 722 TGGGTGGTCCCTCAAGTGAAGTAGAAGACCCGTCAGCCAGACAAAGCAAGTGAAGAA 781
DB |||||
QY 721 TGGGTGGTCCCTCAAGTGAAGTAGAAGACCCGTCAGCCAGACAAAGCAAGTGAAGAA 780
DB |||||
QY 782 GAATCTTCAATAAATGACTACTGAAATGGAATAGAAATTTGATCCCATGCTGGATGAG 841
DB |||||
QY 781 GAATCTTCAATAAATGACTACTGAAATGGAATAGAAATTTGATCCCATGCTGGATGAG 840
DB |||||
QY 842 AGAGGTATTGTTGTTATTTACTGCCGTCGAGGCAACCGCTATTTCGCCGCCGCTCTGTGA 901
DB |||||
QY 841 AGAGGTATTGTTGTTATTTACTGCCGTCGAGGCAACCGCTATTTCGCCGCCGCTCTGTGA 900
DB |||||
QY 902 CTTTACTAGGCTACTACCATATCCATATCTGCTACCAAGGAGGAGAGTCTGTGCT 961
DB |||||
QY 901 CTTTACTAGGCTACTACCATATCCATATCTGCTACCAAGGAGGAGAGTCTGTGCT 960
DB |||||
QY 962 GTCATCTGCTTGTAACTGCTGGTGGCGCGCATCTGCTGGGAGGGTCTAATAGGAGTT 1021
DB |||||
QY 961 GTCATCTGCTTGTAACTGCTGGTGGCGCGCATCTGCTGGGAGGGTCTAATAGGAGTT 1020
DB |||||
QY 1022 TGAGCTCAAAATGCTTAACTGCTGGCAACATATAATAAATGCAATGCTTAAATGATTT 1081
DB |||||
QY 1021 TGAGCTCAAAATGCTTAACTGCTGGCAACATATAATAAATGCAATGCTTAAATGATTT 1080
DB |||||

QY 1082 CTGCTATGAGGCATCTGCCCTCGGTAGCCAGCTCTCCAGAAATTACTTGTAGGTAATTC 1141
DB |||||
QY 1081 CTGCTATGAGGCATCTGCCCTCGGTAGCCAGCTCTCCAGAAATTACTTGTAGGTAATTC 1140
DB |||||
QY 1142 CTCTCTTCATGTTCTAAATAAACTTCTACATTAATACCA 1179
DB |||||
QY 1141 CTCTCTTCATGTTCTAAATAAACTTCTACATTAATACCA 1178
DB |||||
RESULT 9
AR338896 1380 bp DNA linear PAT 17-AUG-2003
LOCUS AR338896
DEFINITION Sequence 387 from patent US 6569662.
ACCESSION AR338896
VERSION AR338896.1 GI:33725753
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1380)
AUTHORS Tang, Y.T., Zhou, P. and Drmanac, R.T.
TITLE Nucleic acids and polypeptides
JOURNAL Patent: US 6569662-A 387 27-MAY-2003;
FEATURES Location/Qualifiers
source 1..1380
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 98.1%; Score 1174; DB 6; Length 1380;
Best Local Similarity 99.6%; Pred. No. 2.1e-280;
Matches 1177; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 CAGCAGTGTCTCTCAGTCTCTCTCAAGCAAGAAAGAGTACTGTGTGAGAGACCAT 60
DB 73 CAGCAGGGTCTCTCAGTCTCTCTCAAGCAAGAAAGAGTACTGTGTGAGAGACCAT 132
QY 61 GGCAGAAATCCTCCAGAGAAATGTGAAGACTGTCACTTCTAAATGCAGAGCTTTTAA 120
DB 133 GGCAGAAATGTGTCCAGAGAAATGTGAAGACTGTCACTTCTAAATGCAGAGCTTTTAA 192
QY 121 ATCCAGAAAAATATGTAAATCACTTAAGATTGTGACCTGGTGTGGTATCTCGGCCCT 180
DB 193 ATCCAGAAAAATATGTAAATCACTTAAGATTGTGACCTGGTGTGGTATCTCGGCCCT 252
QY 181 AACTCTAATTTGCTCTGTTTGGGGAGCAAGCACTTTCTGGCCGGAGGTACCCAAAAAGC 240
DB 253 AACTCTAATTTGCTCTGTTTGGGGAGCAAGCACTTTCTGGCCGGAGGTACCCAAAAAGC 312
QY 241 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAGAGAAATTTACATGGAAT 300
DB 313 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAGAGAAATTTACATGGAAT 372
QY 301 TGATCTGTGACCAAGCACTGAAATATTTCAGAGCGGAAATGGCACTGATGAAACATTTGA 360
DB 373 TGATCTGTGACCAAGCACTGAAATATTTCAGAGCGGAAATGGCACTGATGAAACATTTGA 432
QY 361 AGTGACGACTTTTAAAAACGGATACACTGGCATCTACTCTGGTGGTCTTCAAAAATGTTT 420
DB 433 AGTACACGACTTTTAAAAACGGATACACTGGCATCTACTCTGGTGGTCTTCAAAAATGTTT 492
QY 421 TATCAAACTCAGATTAAGTGATTTCTGAATTTTCTGACCAAGAGAGAGAAATAGATGA 480
DB 493 TATCAAACTCAGATTAAGTGATTTCTGAATTTTCTGAAACCAAGAGAGAGAAATAGATGA 552
QY 481 GAATGAAGAAATTAACCAACTTTCTTTGAACAGTCAAGTCAAGTCAAGTCAAGTCAAGT 540
DB 553 GAATGAAGAAATTAACCAACTTTCTTTGAACAGTCAAGTCAAGTCAAGTCAAGTCAAGT 612
QY 541 GCCTATTGAAAAACCGAGATTTTCTTAAAAATTCAAAAATTCGAGATTTGTGATAACGT 600
DB 613 GCCTATTGAAAAACCGAGATTTTCTTAAAAATTCAAAAATTCGAGATTTGTGATAACGT 672

Dn	860	TGTTGATTTACTGCGTCGAGCAACCGCTATTTCGCCCGCGTCTGTGAACCTTTACTA	919
Qy	911	GGCTACTACCATATCCATCTGCTACCAAGAGACGAGTCATCTGCGTGTCTCATG	970
Dn	920	GGCTACTACCATATCCATCTGCTACCAAGAGACGAGTCATCTGCGTGTCTCATG	979
Qy	971	CTTCTTAACCTGCTGCGCATCTGCGGAGGCTTAATAGGAGGTTTGAGCTCAA	1030
Dn	980	CTTCTTAACCTGCTGCGCATCTGCGGAGGCTTAATAGGAGGTTTGAGCTCAA	1039
Qy	1031	ATGCTTAAACCTGCTGCGCATCTGCGGAGGCTTAATAGGAGGTTTGAGCTCAA	1090
Dn	1040	ATGCTTAAACCTGCTGCGCATCTGCGGAGGCTTAATAGGAGGTTTGAGCTCAA	1098
Qy	1091	AGGATCTGCGGCTGCTGAGCAGCTCTCCAGATTTACTTTAGTAACTCTCTTCA	1150
Dn	1099	AGGATCTGCGGCTGCTGAGCAGCTCTCCAGATTTACTTTAGTAACTCTCTTCA	1158
Qy	1151	TGTTCTAATAAATCTTCACTATTACCAAAAAAATTTTTTTTTTTTTTTTTTTTT	1197
Dn	1159	TGTTCTAATAAATCTTCACTATTACCAAAAAAATTTTTTTTTTTTTTTTTTTTT	1205
RESULT 11			
AF191770			
LOCUS	Homo sapiens myodulin mRNA, complete cds.	1309 bp	mRNA linear PRI 01-NOV-2000
DEFINITION			
ACCESSION	AF191770		
VERSION	AF191770.1	GI:11065923	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 1309)		
AUTHORS	Cros,N., Tkatchenko,A.V., Leclerc,L., Leger,J.J., Marini,J.-F. and Dechesne,C.A.		
TITLE	Gene expression alterations revealed by suppression subtractive hybridization in rat soleus muscle disuse atrophy		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1309)		
AUTHORS	Leclerc,L., Cros,N. and Dechesne,C.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-OCT-1999) Faculte des Sciences, CNRS UMR 6548, Parc Valrose, Nice 06108, France		
FEATURES	Location/Qualifiers		
source	1..1309		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/tissue_type="leg skeletal muscle"		
	187..1140		
	/note="similar to chondromodulin-I"		
	/codon_start=1		
	/product="myodulin"		
	/protein_id="AAG28395.1"		
	/db_xref="GI:11065924"		
	/translation="MAKNPPENCEDCHLNAAEFKSKIKSLKICGLVFLTLTLI		
	VLFWGSHFWEVPEKAYDMHTFYSSGEKKIYMEIDPVTRTEIFBSGNGTDLEV		
	HDKNGTYGIVFGQKCFIKTIQIKVIPFSEPEEEDENEIITTFEQSVIWPAB		
	KPIENRDFLKNKSKILECDNVTMYNPTLISVSELQDFEEBGLDHPANEKKGIEQ		
	NEQWVPOVKYKTRHARQASEELPINDYTENGIEFDPMLDERGYCCIIYCRGNRYC		
	RRVCEPLGGYYPYPCYQGGRVICRVIMPNCNWMVRLGRV"		
CDS			
ORIGIN			
Query Match	98.0%;	Score 1173;	DB 9; Length 1309;
Best Local Similarity	99.6%;	Pred. No. 3.8e-280;	
Matches 1176;	Conservative 0;	Mismatches 5;	Indels 0; Gaps 0;
Qy	1	CAGCAGTGGTCTCTCAGTCCTCTCAAGCAAGGAAGAGTACTGTGTCTGAGAGACCAT	60
Dn	129	CAGCAGTGGTCTCTCAGTCCTCTCAAGCAAGGAAGAGTACTGTGTCTGAGAGACCAT	188

Qy	61	GGCAAGAATCCTCCAGAGAAATTTGTAAGACTGTCAATTTCTAAATGCAGAACTTTTAA	120
Dn	189	GGCAAGAATCCTCCAGAGAAATTTGTAAGACTGTCAATTTCTAAATGCAGAACTTTTAA	248
Qy	121	ATCCAAGAAAATATGTAATCACTTAAAGATTTGTGACTGTGTGTTTGGTATCCTGCCCT	180
Dn	249	ATCCAAGAAAATATGTAATCACTTAAAGATTTGTGACTGTGTGTTTGGTATCCTGCCCT	308
Qy	181	AACTCTAATTTGCTGTTTGGGGAGCAAGCACTTCTGGCCGGAGGTATCCCAAAAAGC	240
Dn	309	AACTCTAATTTGCTGTTTGGGGAGCAAGCACTTCTGGCCGGAGGTATCCCAAAAAGC	368
Qy	241	CTATGACATGGACACATTTCTACAGCAATGGAGAGAAGAAGATTTACATGGAAT	300
Dn	369	CTATGACATGGACACATTTCTACAGCAATGGAGAGAAGAAGATTTACATGGAAT	428
Qy	301	TGATCTCTGTGACCAAGCACTGAAATATTTAGAAAGCGGAAATGGCACTGATGAACATGGA	360
Dn	429	TGATCTCTGTGACCAAGCACTGAAATATTTAGAAAGCGGAAATGGCACTGATGAACATGGA	488
Qy	361	AGTGACACGACTTTTAAACCGGATACACTGGCATCTACTTCTGGGTCTTCAAAAATGTTT	420
Dn	489	AGTGACACGACTTTTAAACCGGATACACTGGCATCTACTTCTGGGTCTTCAAAAATGTTT	548
Qy	421	TATCAAACTCAGATTTAAAGTGATTCCTCAATTTTCTGAACCCAGAGAGAAATAGATGA	480
Dn	549	TATCAAACTCAGATTTAAAGTGATTCCTCAATTTTCTGAACCCAGAGAGAAATAGATGA	608
Qy	481	GAATGAAGAAATATACCAACACTTTTCTTGAACAGTCAGTGATTTGGGTCTCCAGCAAGAAA	540
Dn	609	GAATGAAGAAATATACCAACACTTTTCTTGAACAGTCAGTGATTTGGGTCTCCAGCAAGAAA	668
Qy	541	GCCTATGAAAACCGGAGATTTTCTTAAATTTCCAAATTTCTGGAGATTTGTGATACGT	600
Dn	669	GCCTATGAAAACCGGAGATTTTCTTAAATTTCCAAATTTCTGGAGATTTGTGATACGT	728
Qy	601	GACATGTAATGATCAATCCCACTCTAATATCAGTTTCTGAGTTTCAAGACTTTTGAGGA	660
Dn	729	GACATGTAATGATCAATCCCACTCTAATATCAGTTTCTGAGTTTCAAGACTTTTGAGGA	788
Qy	661	GGAGGAGAGAGATCTTCACTTTCTGCAACGAAAAAAGGAGATTTGAAACAAATGAACA	720
Dn	789	GGAGGAGAGAGATCTTCACTTTCTGCAACGAAAAAAGGAGATTTGAAACAAATGAACA	848
Qy	721	GTGGTGGTCCCTCAAGTGAAAGTAGAAGACCCCTGACCCAGACAAGCAAGTGAAGGA	780
Dn	849	GTGGTGGTCCCTCAAGTGAAAGTAGAAGACCCCTGACCCAGACAAGCAAGTGAAGGA	908
Qy	781	AGAACTTCCCAATAAATGACTATCTGAAATGGAATAGAAATTTGATCCCATGCTGGATGA	840
Dn	909	AGAACTTCCCAATAAATGACTATCTGAAATGGAATAGAAATTTGATCCCATGCTGGATGA	968
Qy	841	GAGAGGTATTGTTGTTTAACTGCGGTGCGGCAACCGCTATTGCGCGCCGCTCTGTGA	900
Dn	969	GAGAGGTATTGTTGTTTAACTGCGGTGCGGCAACCGCTATTGCGCGCCGCTCTGTGA	1028
Qy	901	ACCTTTACTAGGCTACTACCAATATCCATATCTGCTACCAAGGAGCAAGTCACTGTGCG	960
Dn	1029	ACCTTTACTAGGCTACTACCAATATCCATATCTGCTACCAAGGAGCAAGTCACTGTGCG	1088
Qy	961	TGTCATCATGCTTGTAACTGCGGTGCGGCAACCGCTATTGCGGGAGGGTCTTAATAGGAGT	1020
Dn	1089	TGTCATCATGCTTGTAACTGCGGTGCGGCAACCGCTATTGCGGGAGGGTCTTAATAGGAGT	1148
Qy	1021	TTGAGCTCAAAATGCTTAACTGCTGGCAACATATAATAATGCAATGCTATTCAATGAAT	1080
Dn	1149	TTGAGCTCAAAATGCTTAACTGCTGGCAACATATAATAATGCAATGCTATTCAATGAAT	1208
Qy	1081	TCTGCCTATGAGGATCTGCGCCCTGCTAGCAGCTCTCCAGAAATTTACTGTAGTAAAT	1140
Dn	1209	TCTGCCTATGAGGATCTGCGCCCTGCTAGCAGCTCTCCAGAAATTTACTGTAGTAAAT	1268
Qy	1141	CCTCTCTTCAATGTTCTAATAAACTTCTACATTTATCACCAAA	1181

```

Db      1269 CCTCTCTTCATGTTCTAATAAACTTCTACATTAACAATAA 1309
|||||
RESULT 12
AF291656      1184 bp      mRNA      linear      PRI 07-DEC-2001
LOCUS      Homo sapiens chondromodulin-1B mRNA, complete cds.
DEFINITION
ACCESSION
AF291656
VERSION
AF291656.1 GI:15077275
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1184)
Brandau, O., Meindl, A., Fassler, R. and Aszodi, A.
A novel gene, tendin, is strongly expressed in tendons and
ligaments and shows high homology with chondromodulin-1
Dev. Dyn. 221 (1), 72-80 (2001)
MEDLINE
21255555
PUBMED
11357195
REFERENCE
2 (bases 1 to 1184)
Brandau, O., Aszodi, A., Meindl, A. and Fassler, R.
Direct Submission
TITLE
Submitted (30-JUL-2000) Experimental Pathology, Lund University,
Solvegatan 25, Lund 22185, Sweden
JOURNAL
Location/Qualifiers
1. .1184
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
73..1026
/notes="CHM-1B"
/codon_start=1
/product="chondromodulin-1B"
/protein_id="AAK83109.1"
/db_xref="GI:15077276"
/translators="MAKNPPENCEDCHILNABAFKSKIKCSILKICGLVFGILALTLI
VLFPSKHPWPEPKYKAYMEHFYSNGEKKIYMEIDPVTRIELFRSNGTDETLV
HDFKNGYTGIFVGLQKCFIKQIVPESEPEEIDENERITTFPEOSVIVWPAE
KPLENDFGLKILEICDNVTWYINPTLISVSELDQFDEEGEDLHFPANEKKGIBQ
NEQWVQVKEVTRHARQASEBELPINDYENGIEFEDPMLDERGYCCIIYCRGRNYC
RRVCEPLLGYPPYCYQGRVLCRVIMPNNWVARMGRV"
ORIGIN
Query Match      97.7%; Score 1170; DB 9; Length 1184;
Best Local Similarity 100.0%; Pred. No. 2.1e-279;
Matches 1170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGAAAGAGTACTGTGTGCTGAGAGACCAT 60
DB      15 CAGCAGTGGTCTCTCAGTCTCTCAAGCAAGAAAGAGTACTGTGTGCTGAGAGACCAT 74
QY      61 GGCAAGAATCCTCCAGAGAATTGTGAAGACTGTGCACATTCTAAATGCAGAAGCTTTTAA 120
DB      75 GGCAAGAATCCTCCAGAGAATTGTGAAGACTGTGCACATTCTAAATGCAGAAGCTTTTAA 134
QY      121 ATCCAGAAAATATGTAATACATTAGANTTTGGACTGGTGGTGGTATCTGGCCCT 180
DB      135 ATCCAGAAAATATGTAATACATTAGANTTTGGACTGGTGGTGGTATCTGGCCCT 194
QY      181 AACTCTAATTTGCTCTGTTTGGGGGACACAGACTTCTGGCCGAGAGTACCACAAAAGC 240
DB      195 AACTCTAATTTGCTCTGTTTGGGGGACACAGACTTCTGGCCGAGAGTACCACAAAAGC 254
QY      241 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAAGAAGATTTCATGAGAAAT 300
DB      255 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAAGAAGATTTCATGAGAAAT 314
QY      301 TGATCTGTGACCGAAGACTGAATATTTCAGAACGGGAAATGGCACTGATGAACATTGGA 360
DB      315 TGATCTGTGACCGAAGACTGAATATTTCAGAACGGGAAATGGCACTGATGAACATTGGA 374

```

```

QY      361 AGTGCACGACTTTAAAAACGATACACTGGCATCTACTTCTGTGGGTCTTCAAAAATGTTT 420
DB      375 AGTGCACGACTTTAAAAACGATACACTGGCATCTACTTCTGTGGGTCTTCAAAAATGTTT 434
QY      421 TATCAAAACTCAGATTAAAGTGAATTCCTGAATTTTCTGAACCAAGAGAGGAAATAGATGA 480
DB      435 TATCAAAACTCAGATTAAAGTGAATTCCTGAATTTTCTGAACCAAGAGAGGAAATAGATGA 494
QY      481 GAATGAAGAAATATACCAAACTTTCTTTGAACAGTCAAGTGAATTTGGGTCCACAGAGAAA 540
DB      495 GAATGAAGAAATATACCAAACTTTCTTTGAACAGTCAAGTGAATTTGGGTCCACAGAGAAA 554
QY      541 GCCTATTGAAAAACCGAGATTTCCTTAAAAATTCCTTAAAAATTCCTGAGAGATTTCGTGATAACGT 600
DB      555 GCCTATTGAAAAACCGAGATTTCCTTAAAAATTCCTTAAAAATTCCTGAGAGATTTCGTGATAACGT 614
QY      601 GACCATGATTGGATCAATCCCACTCTTAATATCAGTTTCTGAGTTACAAGACTTTTGAGGA 660
DB      615 GACCATGATTGGATCAATCCCACTCTTAATATCAGTTTCTGAGTTACAAGACTTTTGAGGA 674
QY      661 GAGGGGAGAAAGATCTTCACCTTCTGCGCAACGAAAAAAGGAGATTGAACAAAAATGAACA 720
DB      675 GAGGGGAGAAAGATCTTCACCTTCTGCGCAACGAAAAAAGGAGATTGAACAAAAATGAACA 734
QY      721 GTGGGTGTCCTCAAGTGAAGTAGAAGACCGGTACGCCAGACCAAGCAAGTGAAGGA 780
DB      735 GTGGGTGTCCTCAAGTGAAGTAGAAGACCGGTACGCCAGACCAAGCAAGTGAAGGA 794
QY      781 AGAACTTCCAAATAATGACTATATCTGAAAAATGAATAAATTTGATCCCATGCTGGAATGA 840
DB      795 AGAACTTCCAAATAATGACTATATCTGAAAAATGAATAAATTTGATCCCATGCTGGAATGA 854
QY      841 GAGAGGTTATTGTGTTATTTACTGCGGTGAGGCAACCGCTATTGCGCGCGCGCTCTGTGA 900
DB      855 GAGAGGTTATTGTGTTATTTACTGCGGTGAGGCAACCGCTATTGCGCGCGCGCTCTGTGA 914
QY      901 ACCTTTACTAGGCTACTACCCTATCCATATCTGCTACCAAGGAGGAGTCAATCTGTGCG 960
DB      915 ACCTTTACTAGGCTACTACCCTATCCATATCTGCTACCAAGGAGGAGTCAATCTGTGCG 974
QY      961 TGTCAATCATGCTTTGTAACCTGTGGGTGGCGCGCATGCTGCGGAGGGTCTTAATAGAGGAGT 1020
DB      975 TGTCAATCATGCTTTGTAACCTGTGGGTGGCGCGCATGCTGCGGAGGGTCTTAATAGAGGAGT 1034
QY      1021 TTGAGCTCAAAATGCTTAAACTGCTGGCAACATATATAAATGATGCTATTCATGAATTT 1080
DB      1035 TTGAGCTCAAAATGCTTAAACTGCTGGCAACATATATAAATGATGCTATTCATGAATTT 1094
QY      1081 TCTGCTCTATGAGGCATCTGGCCCTCTGTAGCCAGCTCTCCAGAAATTAATTTAGTGAATTT 1140
DB      1095 TCTGCTCTATGAGGCATCTGGCCCTCTGTAGCCAGCTCTCCAGAAATTAATTTAGTGAATTT 1154
QY      1141 CCTCTCTTCATGTTCTTAATAAACTTCTTACA 1170
DB      1155 CCTCTCTTCATGTTCTTAATAAACTTCTTACA 1184
RESULT 13
AB055421
LOCUS      Homo sapiens mRNA for CHML, complete cds.
DEFINITION
ACCESSION
AB055421
VERSION
AB055421.1 GI:12698292
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Yamana, K., Wada, H., Takahashi, Y., Sato, H., Kasahara, Y. and
Kiyoki, M.
Molecular cloning and characterization of CHML, a novel membrane
TITLE

```

molecule similar to chondromodulin-I		Biochem. Biophys. Res. Commun. 280 (4), 1101-1106 (2001)	
JOURNAL	21092728		
MEDLINE	11162640		
PUBMED	2 (bases 1 to 1140)		
REFERENCE	Yamana,K.		
AUTHORS	Direct Submission		
TITLE	Submitted (05-FEB-2001) Kei Yamana, Teijin Limited, Teijin		
JOURNAL	Institute for Biomedical Research; 4-3-2 Asahigaoka, Hino, Tokyo		
	191-8512, Japan (E-mail:k.yamana@teijin.co.jp,		
	Tel.81-42-586-8281(ex.8281), Fax:81-42-587-5519)		
FEATURES	Location/Qualifiers		
source	1..1140		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
gene	/db_xref="taxon:9606"		
	1..1140		
	/gene="ChM1L"		
CDS	7..960		
	/gene="ChM1L"		
	/codon_start=1		
	/product="ChM1L"		
	/protein_id="BAB21756.1"		
	/db_xref="GI:12698293"		
	translation="WAKNPNCEDCHIILNABAFKSKIKSLKICGLVFGILALTLI		
	VLFWSGKHFPWEPKADYMEHTFYSGEKKKIYMEIDVETREIFRSGNGDTLEV		
	HDFKNGYTGIVFGLQKCFIKTKIVPESEPEEIDENEIITTFPFSQSVIWPAAE		
	KPIENRDLFKSLKILEICDNVTWYNPTLISVSELODPREBEGDLHPANKEKGLIQ		
	NEQWVPOVKTEKTHARQASBEELPINDYTENGIEFDPMLEDERGYCCIIYCRNGRYC		
	RVCPPELLGYFPYCYQSGGRVICRVIMPCNWWVARMGRV"		
ORIGIN			
Query Match	95.1%;	Score 1138.4;	DB 9; Length 1140;
Best Local Similarity	99.9%;	Pred. No. 1.5e-271;	
Matches 1139;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
QY	53	GAGACATGCGAAGAAATCCTCCAGAGAATTGTGAAGACTGTGCACATTTCAATGCGAGAA	112
DB	1	GAGACATGCGAAGAAATCCTCCAGAGAATTGTGAAGACTGTGCACATTTCAATGCGAGAA	60
QY	113	GCCTTTTAAATCCAGAAAATATGTAATCACTTAAGATTGTGACTGGTGTGTGTATC	172
DB	61	GCCTTTTAAATCCAGAAAATATGTAATCACTTAAGATTGTGACTGGTGTGTGTATC	120
QY	173	CTGGCCCTAACTCTAATCTCTCTGTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCC	232
DB	121	CTGGCCCTAACTCTAATCTCTCTGTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCC	180
QY	233	AAAAAGCCTATGACATGAGCAGACACTTCTACAGCAATGGAGAGAGAGATTAC	292
DB	181	AAAAAGCCTATGACATGAGCAGACACTTCTACAGCAATGGAGAGAGAGATTAC	240
QY	293	ATGGAATTTGATCCTGTGACCAAGCACTGAAATATTCAGAAAGCGAAATGGCACTGATGAA	352
DB	241	ATGGAATTTGATCCTGTGACCAAGCACTGAAATATTCAGAAAGCGAAATGGCACTGATGAA	300
QY	353	ACATTGGAAGTGCAGACTTTTAAAAACCGATACACTGGCATCTACTTCGTGGGTCTTCAA	412
DB	301	ACATTGGAAGTGCAGACTTTTAAAAACCGATACACTGGCATCTACTTCGTGGGTCTTCAA	360
QY	413	AAATGTTTTATCAAACTCAGATTAAAGTGATCTCGAATTTCTGAACAGAGAGAA	472
DB	361	AAATGTTTTATCAAACTCAGATTAAAGTGATCTCGAATTTCTGAACAGAGAGAGAA	420
QY	473	ATAGATGAGNATGAAGAAATTTACCAACTTTCTTTTGAACAGTCACTGATTTGGGTCCCA	532
DB	421	ATAGATGAGNATGAAGAAATTTACCAACTTTCTTTTGAACAGTCACTGATTTGGGTCCCA	480
QY	533	GCAGAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCGAAAAATTCGAGAGATTGT	592
DB	481	GCAGAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCGAAAAATTCGAGAGATTGT	540
QY	593	GATAACGTGACCATGATTGATGATCAATCCCACTCTAATATACATTTCTGAGTTTACAGAC	652
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			

Db 193 ATCCAAGAAAATATGTAAATCATTTAAGATTGTGGACTGGTGTGGTATCTCTGGCCCT 252

Qy 181 AACTCTAATGTCCTGTTTGGGGAGCAGACATTTGCGCCGGAGGTACCCAAAAAGC 240

Db 253 AACTCTAATGTCCTGTTTGGGGAGCAGACATTTGCGCCGGAGGTACCCAAAAAGC 312

Qy 241 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAGAAAGATTTCATCGAAT 300

Db 313 CTATGACATGGAGCACACTTTCTACAGCAATGGAGAGAGAAAGATTTCATCGAAT 372

Qy 301 TGATCTGTGACCGAATCTGAATAATTGAGAGCGGAATGGCACTGATGAACATTGGA 360

Db 373 TGATCTGTGACCGAATCTGAATAATTGAGAGCGGAATGGCACTGATGAACATTGGA 432

Qy 361 AGTCACGACCTTAAACACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTT 420

Db 433 AGTCACGACCTTAAACACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTT 492

Qy 421 TATCAAAACTCAGATTAAAGTGAATCCTGAAATTTCTGAACAGAGAGAAATAGATGA 480

Db 493 TATCAAAACTCAGATTAAAGTGAATCCTGAAATTTCTGAACAGAGAGAAATAGATGA 552

Qy 481 GAATGAAGAAATACCACAACTTTCTTTGAACAGTCAGTGATTTGGGTCCCGAGCAGAAA 540

Db 553 GAATGAAGAAATACCACAACTTTCTTTGAACAGTCAGTGATTTGGGTCCCGAGCAGAAA 612

Qy 541 GCCTATTGAACACCGAGATTTCCTTAAAAATCCAAAATCTGGAGATTGTGATAAGT 600

Db 613 GCCTATTGAACACCGAGATTTCCTTAAAAATCCAAAATCTGGAGATTGTGATAAGT 672

Qy 601 GACCATGTATTGGATCAATCCACACTTAATATC----- 633

Db 673 GACCATGTATTGGATCAATCCACACTTAATATCAGGAACAATTCGAAGCAGTTGCATCA 732

Qy 634 -----AGTTTCTGAGTTCAAGACTTTTGAGGAGGGGAGAGA 672

Db 733 CAACCTTTGCAATTTATCTTAGTTTCTGAGTTTCAAGACTTTTGAGGAGGGGAGAGA 792

Qy 673 TCTTCACCTTCCTGCCACGAAAGAAAGGATTGAACAAATGACAGTGGTGGTCC 732

Db 793 TCTTCACCTTCCTGCCACGAAAGAAAGGATTGAACAAATGACAGTGGTGGTCC 852

Qy 733 TCAAGTGAAGTAGAGAAGACCCGTCAGCCAGACAAAGTGAAGAGAACTTCCAA 792

Db 853 TCAAGTGAAGTAGAGAAGACCCGTCAGCCAGACAAAGTGAAGAGAACTTCCAA 912

Qy 793 AAATGACTATCTGAAAATGGAATAGAAATTTGATCCCATGCTGGATGAGAGAGTTATG 852

Db 913 AAATGACTATCTGAAAATGGAATAGAAATTTGATCCCATGCTGGATGAGAGAGTTATG 972

Qy 853 TTGTATTACTGCGCTGAGGCAACCGTATTGCGCGCGCTGTGTGAACCTTTACTAGG 912

Db 973 TTGTATTACTGCGCTGAGGCAACCGTATTGCGCGCGCTGTGTGAACCTTTACTAGG 1032

Qy 913 CTACTACCATATCATACTGCTACCAAGGAGGAGCAGTCACTGCTGCTCATCATGCC 972

Db 1033 CTACTACCATATCATACTGCTACCAAGGAGGAGCAGTCACTGCTGCTCATCATGCC 1092

Qy 973 TTGTAACTGGTGGGTCGCCATGCTGGGAGGAGTCTAAATAGGAGGTTTGAGCTCAAT 1032

Db 1093 TTGTAACTGGTGGGTCGCCATGCTGGGAGGAGTCTAAATAGGAGGTTTGAGCTCAAT 1152

Qy 1033 GCTTAACTGCTGGCAACATATAAATGATGCTATTCATGAATTTCTGCCTATGAG 1092

Db 1153 GCTTAACTGCTGGCAACATATAAATGATGCTATTCATGAATTTCTGCCTATGAG 1212

Qy 1093 GCATCTGCGCCCTGGTAGCAGCTCTCCAGAAATTAATGTTAGGTAATTCCTCTTCATG 1152

Db 1213 GCATCTGCGCCCTGGTAGCAGCTCTCCAGAAATTAATGTTAGGTAATTCCTCTTCATG 1272

Qy 1153 TTCTAATAAACTTCTACATTTATCAACAAA 1182

Db 1273 TTCTAATAAACTTCTACATTTATCACAACA 1302

RESULT 15

AB059407

LOCUS

DEFINITION

AB059407

ACCESSION

AB059407.2

VERSION

GI:23200576

KEYWORDS

SOURCE

ORGANISM

Equus caballus (horse)

Equus caballus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

REFERENCE

1

Hasegawa,T.

Molecular cloning and characterization of mRNA for equine tenomodulin

Published Only in Database (2001)

2

(bases 1 to 1175)

Hasegawa,T.

Direct Submission

Submitted (05-APR-2001) Telhisa Hasegawa, JRA Equine Research Institute, Laboratory of Molecular and Cellular Biology; 321-4 Tokami-Cho, Utsunomiya, Tochigi 320-0856, Japan (E-mail:telhisah@center.equinist.go.jp, Tel:81-28-647-0662, Fax:81-28-647-0686)

On Sep 19, 2002 this sequence version replaced gi:15982572.

COMMENTS

Location/Qualifiers

1..1175

/organism="Equus caballus"

/mol_type="mRNA"

/strain="Thorougbred"

/db_xref="taxon:9796"

/tissue_type="tendon"

/dev_stage="adult"

/note="domestic horse"

1..1175

/gene="TNMD"

52..1005

/gene="TNMD"

/codon_start=1

/product="tenomodulin"

/protein_id="BAB69482.2"

/db_xref="GI:23200577"

/translation="MAKNPPENCEDCHILNAEAFKSKICKSLKICGLVFGILALTILVLFWGGHFPETPKTYDMEHTYSNGEKKIYMEIDPVTRTIFRSNGTDETLVHDFKNGYTIYFVGLQKCFIKIQKVPFSEPEEEDNEEITTFPEQSVIWPVAKPIENRDLKNSKILEICDNVTWYNPTLILAVSELODFEEDGDLHPTNDKKGIEONEQVVPQVKVETRRARQASEELPINDYTENGIEFDPMLDERGYCCICYCRGNRYCRRVCEPLIGYPPYCYQGGRVICRVIMPCNWWVARMGLRV"

polyA_signal

1154..1159

/gene="TNMD"

1175

/gene="TNMD"

ORIGIN

Query Match 87.2%; Score 1043.6; DB 4; Length 1175;

Best Local Similarity 93.8%; Pred. No. 4.9e-248;

Matches 1099; Conservative 0; Mismatches 69; Indels 4; Gaps 1;

Qy 8 GGTCTCTCAGTCTCTCAAGCAAGGAAGAGTACTGTGCTGAGAGACCATGCCAAG 67

Db 1 GGTCTCTCAGTCTCTCAAGCAAGGAAGAGTACTGTGCTGAGAGACCATGCCAAG 60

Qy 68 AATCCTCCAGAAATTTGAAAGCTGTCAATTCATAATGCAGAACTTTTAAATCCAAG 127

Db 61 AATCCTCCAGAAATTTGAAAGCTGTGAGACTGTCAATTTTAAATGCAGAACTTTTAAATCCAAG 120

Qy 128 AAAATATGTAATCACTAAAGATTTGTGGACTGTGTGTTGTGTTCTCTGGCCTAACTCTA 187

Db 121 AAGATATGTAATCACTAAATTTGTGGATTTGTGTTGTGTTCTCTGGCCTAACTCTA 180

QY 188 ATTGTCCTGTTTGGGGAGCAAGCACTTCTGCGGAGGTACCCAAAAAGCCTATGAC 247
Db |||||
181 ATTGTCCTGTTTGGGGAGCAAGCACTTCTGCGGAGGTACCCAAAAAGCCTATGAC 240
QY 248 ATGGAGCACACTTCTACAGCAATGAGAGAAAGAGATTTACATGGAAATGATCCT 307
Db |||||
241 ATGGAGCACACTTCTACAGCAACGGAGAGAGAGAAATTTACATGGAAATGATCCC 300
QY 308 GTGACCAGAACTGAAATATTTCAGAACGGGAAATGGCACTGATGAAACATTTGGAAGTGCAC 367
Db |||||
301 GTGACCAGAACTGAAATATTTCAGAACGGGAAATGGCACTGATGAAACATTTGGAAGTGCAC 360
QY 368 GACTTTAAAAACCGATACACTGCGACTTCTCGTGGGTCTTCAAAAATGTTTTATCAAA 427
Db |||||
361 GACTTTAAAAATGGATATACTGCGCATCTACTTTGTAGGTCTTCAAAAATGCTTCAACAA 420
QY 428 ACTCAGATTAAAGTGATTCTGTGAATTTCTGACCGAGAGGAAATAGATGAGAAATGAA 487
Db |||||
421 ACTCAGATTAAAGTGATTCTGTGAATTTCTGACCGAGAGGAAATAGATGAGAAATGAA 480
QY 488 GAAATTACCAACAACCTTCTTTGAACAGTCACTGATTTGGGTCCAGCAGAAAAAGCCTATT 547
Db |||||
481 GAAATTACCAACAACCTTCTTTGAACATCACTGATTTGGGTCCAGCAGAAAAAGCCTATT 540
QY 548 GAAAAACGAGATTTCTTAAAAAATTCAAAAATTTCTGGAGATTTGTGATAACGTGACCATG 607
Db |||||
541 GAAAAACGAGACTTCTTAAAAAATTCAAAAATTTCTGGAGATTTGTGATAACGTGACCATG 600
QY 608 TATTGGATCAATCCCACTCTAAATATCAGTTTCTGAGTTTACAAGACTTTGAGGAGGAGGA 667
Db |||||
601 TATTGGATCAATCCCACTCTAAATAGAGTTTCAAGACTTTGAGGAGGATGGT 660
QY 668 GAAGATCTTCACTTCTGCGCAACGAAAAAAGGATTTGAACAAAAATGAACACAGTGGGTG 727
Db |||||
661 GAAGATCTTCACTTCTGCGCAACGAAAAAAGGATTTGAACAAAAATGAACACAGTGGGTG 720
QY 728 GTCCCTCAAGTGAAGTAGAGAGACCCCGTCAAGCAGACAGCAAGCAAGTGAAGAACTT 787
Db |||||
721 GTCCCTCAAGTGAAGTGAAGAGACCCCGTCAAGCAGACAGCAAGCAAGTGAAGAACTT 780
QY 788 CCAATAATGACTATCTGAAATGGAATAGATTTGATCCCATGCTGGATGAGAGAGT 847
Db |||||
781 CCAATAATGACTATCTGAAATGGAATAGATTTGATCCCATGCTGGATGAGAGAGT 840
QY 848 TATTGTTGTATTTACTGCCGTGAGGCAACCGTATTGCGCGCGCGTCTGTGAACCTTTA 907
Db |||||
841 TATTGTTGTATTTACTGCCGTGAGGCAACCGTATTGCGCGCGCGTCTGTGAACCTTTA 900
QY 908 CTAGGCTACTACCCATATCCATCTGCTACAGGAGGACGAGTCATCTGTCGTGTCATC 967
Db |||||
901 CTAGGTTACTACCCGTATCCATCTGCTACCAAGGAGGGGGTTATCTGTCGTGTCATC 960
QY 968 ATGCTTTGTAACGTGGTGGCGCGCATGCTGGGAGGGGTCTAATAGGAGGTTTGAGCT 1027
Db |||||
961 ATGCTTTGCAACTGGTGGTGGCGCGCATGTTGGGGAGGGTCTAGTAGCAAGTTCGAGCA 1020
QY 1028 CAAATGCTTAAACTGTGSG---CAACATATAATAAATGCAATGCTATTCATGAATTTCT 1083
Db |||||
1021 CAAGAGCTTAAATTTCTGGCAGCAACATATAATAAATGATGCTATTCATGAATTTCT 1080
QY 1084 GCCTATGAGGATCTGCGCCCTGGTAGCCAGCTCTCCAGAAATTAATTTGTTAGGTAATTCCT 1143
Db |||||
1081 CCCTATAAGGCAATTTGGCTCTCGTAGCCAGTACTCCAGAAATTAATTTGTTAGGTAATTCCT 1140
QY 1144 CTCCTCATGTTCTAATAAATTTCTACATATC 1175
Db |||||
1141 CTCCTCATGTTCTAATAAATTTCTACATATC 1172